

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 12, 2006, 06:25:40 ; Search time 19961 Seconds  
(without alignments)  
20089.779 Million cell updates/sec

Title: US-09-522-753-4  
Perfect score: 8564.8  
Sequence: 1 catgctcgagcctccacacagc.....caaaaaaaaaaaaaaaaa 8571

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_est3:  
4: gb\_est4:  
5: gb\_est5:  
6: gb\_est6:  
7: gb\_est7:  
8: gb\_est8:  
9: gb\_est9:  
10: gb\_est10:  
11: gb\_est11:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5140.2	60.0	7372	10 AY412686	AY412686 Homo sapi
2	4216.6	49.2	6836	10 AY412687	AY412687 Pan trogl
3	3315.2	38.7	7013	10 AY412688	AY412688 Mus muscu
4	1473.6	17.2	2000	4 BC030087	BC030087 Homo sapi
5	1297.4	15.1	1362	4 BC020427	BC020427 Homo sapi
6	887.8	10.4	1093	3 BM909096	BM909096 AGENCOURT
7	868.6	10.1	965	3 BM555371	BM555371 AGENCOURT
8	841.8	9.8	846	5 BU557144	BU557144 AGENCOURT
9	815.6	9.5	1126	3 BM553310	BM553310 AGENCOURT
10	798.2	9.3	984	5 BQ690869	BQ690869 AGENCOURT
11	796.4	9.3	1051	3 BM423558	BM423558 AGENCOURT
12	796.2	9.3	1159	3 BM802749	BM802749 AGENCOURT
13	795.2	9.3	1057	3 BM477568	BM477568 AGENCOURT
14	787.6	9.2	1022	3 BM910785	BM910785 AGENCOURT
15	782	9.1	875	5 BQ711119	BQ711119 AGENCOURT
16	779	8.9	1010	3 BM915686	BM915686 AGENCOURT
17	766.2	8.9	1030	3 BM558844	BM558844 AGENCOURT
18	760.8	8.9	882	5 BX397973	BX397973 AGENCOURT
19	754.6	8.8	892	3 BU538827	BU538827 AGENCOURT
20	748.8	8.7	1036	3 BM910476	BM910476 AGENCOURT
21	747.6	8.7	868	5 BX368971	BX368971 AGENCOURT
22	746.8	8.7	882	5 BU180236	BU180236 AGENCOURT

23	746	8.7	923	5 BX390462	BX390462 AGENCOURT
24	743.4	8.7	1075	3 BM461469	BM461469 AGENCOURT
25	740	8.6	916	5 BQ892847	BQ892847 AGENCOURT
26	737.4	8.6	1045	3 BM560255	BM560255 AGENCOURT
27	730.8	8.5	1038	3 BM910704	BM910704 AGENCOURT
28	730.6	8.5	1066	3 BM471347	BM471347 AGENCOURT
29	728.4	8.5	882	2 BI089430	BI089430 AGENCOURT
30	720.8	8.4	769	5 BX368972	BX368972 AGENCOURT
31	712.6	8.3	958	5 BU172348	BU172348 AGENCOURT
32	711.8	8.3	958	2 BG831424	BG831424 AGENCOURT
33	711.4	8.3	880	8 CV806970	CV806970 AGENCOURT
34	711	8.3	875	5 BQ691710	BQ691710 AGENCOURT
35	707.4	8.3	923	5 BU184403	BU184403 AGENCOURT
36	706.8	8.3	1094	3 BM560912	BM560912 AGENCOURT
37	703	8.2	876	2 BG252161	BG252161 AGENCOURT
38	703	8.2	1067	3 BM472005	BM472005 AGENCOURT
39	700.6	8.2	928	5 BU164114	BU164114 AGENCOURT
40	695.8	8.1	772	6 CF135847	CF135847 AGENCOURT
41	692.8	8.1	727	5 BU632778	BU632778 AGENCOURT
42	692	8.1	830	7 CO648233	CO648233 AGENCOURT
43	688	8.0	861	5 BU191146	BU191146 AGENCOURT
44	683	8.0	737	6 CD742837	CD742837 AGENCOURT
45	679.6	7.9	1087	2 BG252257	BG252257 AGENCOURT

## ALIGNMENTS

RESULT 1  
AY412686  
LOCUS  
DEFINITION  
Homo sapiens NCOR2 gene, VIRTUAL TRANSCRIPT, partial sequence.  
ACCESSION  
AY412686  
VERSION  
GI:39768651  
KEYWORDS  
GSS.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE  
1 (bases 1 to 7372)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,  
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
Adams,M.D. and Cargill,M.  
Interfering nonneutral evolution from human-chimp-mouse orthologous  
gene ctios  
JOURNAL  
Science 302 (5652), 1960-1963 (2003)  
PUBMED  
14671302  
AUTHORS  
2 (bases 1 to 7372)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,  
Fertiera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
Adams,M.D. and Cargill,M.  
Direct Submission  
JOURNAL  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source  
1..7372  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..>7372  
/gene="NCOR2"  
/locus\_tag="HOM4640"  
ORIGIN  
Query Match : 60.0%; Score 5140.2; DB 10; Length 7372;  
Best Local Similarity 69.9%; Pred. No. 0;  
Matches 5151; Conservative 0; Mismatches 2221; Indels 0; Gaps 0;

OY		2	AATGTGGGCTCCACAAGCTTGTGGCAGAAGTGTGAAGGCCACTGAACCCCGCTAACCG	61
Db		1	AATGTGGGATCCACAAGCTGTGGCAAGA GTGAAGGCCACTGAACCCCGCTAACCG	60
OY		62	CCCCAAGCTTTTCTTAACCAAGTAGAGTCGCCTGGAGCAACA CGAAGTGCGGCTCTTG	121
Db		61	CCCCAAGCTTTTCTTAACCAAGTAGAGTCGCCTGGAGCAACA CGAAGTGCGGCTCTTG	120
OY		122	GAGTACCAAGACCACTCCCAGCATATAGCTCTCCACTCTGTGCGCGGGCTCATATCCAG	181
Db		121	GAGTACCAAGACCACTCCCAGCATATAGCTCTCCACTCTGTGCGCGGGCTCATATCCAG	180
OY		182	CCCCAGCGCGGAGGCGCTCCCTGCTGTGTGAGTTCCAGCCCGGAAATGAACGGTCCAG	241
Db		181	CCCCAGCGCGGAGGCGCTCCCTGCTGTGTGAGTTCCAGCCCGGAAATGAACGGTCCAG	240
OY		242	GAGCTTCAACTGCGGCGAAGTCCCACTCATACCTGCCCCAGCTGGGAACTCAGAATG	301
Db		241	NNN	300
OY		302	GAGTTCATTGAAGAAGAGCGCCCTGGCTAGAGCTGTGCTGACCCCTGTCTGACCG	361
Db		301	NNN	360
OY		362	TCACCCTGTCTGGCCACGCGGCGAGCTCGGATCTGAAGACTTCAACAAGACCGTAGC	421
Db		361	NNN	420
OY		422	CTGACGGGCAAGCTTGAACCGGTGTCTCCCCAGCCCCCGCACACTGACCCTGAAGCTG	481
Db		421	CTGACGGGCAAGCTTGAACCGGTGTCTCCCCAGCCCCCGCACACTGACCCTGAAGCTG	480
OY		482	GAGCTGTGTGCGGCACCGGCTGTCCAAGAGAGAGCTGATCCAGAACTGAACCGGCTGAC	541
Db		481	GAGCTGTGTGCGGCACCGGCTGTCCAAGAGAGAGCTGATCCAGAACTGAACCGGCTGAC	540
OY		542	CGAAGATTCACCATGTGTAGAGACGAGATCTTTAAAGCTGAAGAAAGACAGCAACGCTG	601
Db		541	CGAAGATTCACCATGTGTAGAGACGAGATCTTTAAAGCTGAAGAAAGACAGCAACGCTG	600
OY		602	GAGGAGGAGGCTGCCAAGCCGCGGAGCTAGAAAGCCCGTGCACGCGCCCATCTGAG	661
Db		601	GAGGAGGAGGCTGCCAAGCCGCGGAGCTAGAAAGCCCGTGCACGCGCCCATCTGAG	660
OY		662	TGGAAGCACCGGAGCCTGTGACATCACTTACGACGAAACCGGAMAAGCTGAAGCT	721
Db		661	TGGAAGCACCGGAGCCTGTGACATCACTTACGACGAAACCGGMAAAGCTGAAGCT	720
OY		722	GCAKATCGGATTTCTGAAGGCTGGGGGCCCAAGGTGTGAGCTGCGCTTACAACAAGCC	781
Db		721	NNN	780
OY		782	TCCGACACCCGGCACTATCATGAGAACTCAAATAAACAAGCGCATCGGAAGAAGCTA	841
Db		781	TCCGACACCCGGCACTATCATGAGAACTCAAATAAACAAGCGCATCGGAAGAAGCTA	840
OY		842	ATCTTTGTACTTCAAGAGAGGAATCAGCTCGGAAAACATGAGAAAGAGATTCTGCCAG	901
Db		841	ATCTTTGTACTTCAAGAGAGGAATCAGCTCGGAAAACATGAGAAAGAGATTCTGCCAG	900
OY		902	CGCTATGACCAAGCTCATGAGAGGCTTTGAAAAAAGGTGTGAGCCGATCGAAAACAACCG	961
Db		901	CGCTATGACCAAGCTCATGAGAGGCTTTGAAAAAAGGTGTGAGCCGATCGAAAACAACCG	960
OY		962	CGCCGCGCGGCGCAAGAGAGCAAGGTGTGCGGAGTACTGAAAAGCAATTTCTCTGAGATC	1021
Db		961	CGCGCGCGGCGCAAGAGAGCAAGGTGTGCGGAGTACTGAAAAGCAATTTCTCTGAGATC	1020
OY		1022	CGCAAGCAAGCGGAGCTGACGAGCGCATGCAAGACAGGCTGTGCGCACGCGGCGAGTGGG	1081
Db		1021	CGCAAGCAAGCGGAGCTGACGAGCGCATGCAAGACAGGCTGTGCGCACGCGGCGAGTGGG	1080

Qy	1082	CTGTCCATGTCCGCGCCGCGAGGAGAGAGAGGTGTCAAGATCATCGATGGCCCTCA	1141
Dp	1081	CTGTCCATGTCCGCGCCGCGAGGAGAGAGAGGTGTCAAGATCATCGATGGCCCTCA	1144
Qy	1142	GAGCAGGAGAACTCGAGAAAGCAGATGGCGCAGCTGGCCGTGATCCGCCATGCTGTAC	1201
Dp	1141	GAGCAGGAGAACTCGAGAAAGCAGATGGCGCAGCTGGCCGTGATCCGCCATGCTGTAC	1200
Qy	1202	GACGCTGACCAAGCAGCGCATCAAGTTCAACAATGAACGGGCTTTATGGCCGACCCATG	1261
Dp	1201	GACGCTGACCAAGCAGCGCATCAAGTTCAACAATGAACGGGCTTTATGGCCGACCCATG	1260
Qy	1262	AAGGTGTACAAAGACCCGCGAGTATGAACAATGTGGATGTGACGAGGAAGAGACCTTC	1321
Dp	1261	AAGGTGTACAAAGACCCGCGAGTATGAACAATGTGGATGTGACGAGGAAGAGACCTTC	1320
Qy	1322	CGGAGGAAGTTCAATGACAGCATCCCAAGAACTTTGGCTGATCGCATCATTCCTGGAGAG	1381
Dp	1321	CGGAGGAAGTTCAATGACAGCATCCCAAGAACTTTGGCTGATCGCATCATTCCTGGAGAG	1380
Qy	1382	AAGCAGTGGCTGAGTGGCTCTCTATTTACTGACTTAAGAAATGAGAACTATTAAG	1441
Dp	1381	NNNACATGGCTGAGTGGCTCTCTATTTACTGACTTAAGAAATGAGAACTATTAAG	1440
Qy	1442	AGCCTGTGTGACGAGCATATCGGCGCGCGGCAAGGCCGACGACCAACAACACAGCAG	1501
Dp	1441	AGCCTGTGTGACGAGCATATCGGCGCGCGGCAAGGCCGACGACCAACAACACAGCAG	1500
Qy	1502	CAGCAGCAGCAGCAGCAGCAGCAGCAGCCATGCCCCGACAGCAGAGAGAGAA	1561
Dp	1501	NN	1560
Qy	1562	GATGAGAGAGAGAGAGAGAGAGCGAGAGAGAGAGAGAGAAACCGGAGGTGAGAAC	1621
Dp	1561	NN	1620
Qy	1622	GACAGAGAGAACTCTCTCAAGAGAGAAACAACAACCTTAAGGAGAGAGAAACAACAAG	1681
Dp	1621	NN	1680
Qy	1682	AAGGAGGCTGTGGCCTCCAAAGGCGCGCAAACTGCAACAAGCCAGGGAAGACGCAAAAGC	1741
Dp	1681	AAGGAGGCTGTGGCCTCCAAAGGCGCGCAAACTGCAACAAGCCAGGGAAGACGCAAAAGC	1740
Qy	1742	CGCATCAACCGGCTCAATGGCTAATGAGGCCAACAAGGAGAGGCCATCACCCCCACAG	1801
Dp	1741	CGCATCAACCGGCTCAATGGCTAATGAGGCCAACAAGGAGAGGCCATCACCCCCACAG	1800
Qy	1802	AGCGCCGAGCTGGCCTCCATGTGAGCTGAATGAAGTTCTCGCTGACCAAGAAAGAAATG	1861
Dp	1801	AGCGCCGAGCTGGCCTCCATGTGAGCTGAATGAAGTTCTCGCTGACCAAGAAAGAAATG	1860
Qy	1862	GAAACAGCCCAAGAAAGGTCTCTGTGAACAACGCGCGCAACTGTGGGCATGCGCCGATG	1921
Dp	1861	GAAACAGCCCAAGAAAGGTCTCTGTGAACAACGCGCGCAACTGTGGGCATGCGCCGATG	1920
Qy	1922	GTGGGCTCCAAAGCTGTGTGCAAGTGTAAAGCTTTCACTTCAACTACAAGAAAGGCGAG	1981
Dp	1921	GTGGGCTCCAAAGCTGTGTGCAAGTGTAAAGCTTTCACTTCAACTACAAGAAAGGCGAG	1980
Qy	1982	AACTCGATGATGATCTTTCAGACAGCAACAAGCTGAAGATGAGAGAGAGAGAAACGCGCG	2041
Dp	1981	AACTCGATGATGATCTTTCAGACAGCAACAAGCTGAAGATGAGAGAGAGAGAAACGCGCG	2040
Qy	2042	AGGAAGAAAGAAAGGCGCGCGCGCGGCAAGCGAGAGAGCTGGATTCGCGCCGTGGTG	2101
Dp	2041	AGGAAGAAAGAAAGGCGCGCGCGCGGCAAGCGAGAGAGCTGGATTCGCGCCGTGGTG	2100
Qy	2102	GAGGATGAGAGATGTGAGGCGTCTGGGCGTGTGACGGAATGTGAGAGAGAGATGTGTGAGAG	2161
Dp	2101	GAGGATGAGAGATGTGAGGCGTCTGGGCGTGTGACGGAATGTGAGAGAGAGATGTGTGAGAG	2160
Qy	2162	GCTGAAGCCTTACATGCGCTCTGGGAATGAGGTGCCAAGGGGAATGATGTGGCCACGC	2221

[illegible][illegible]

[illegible]

QY	5462	ATGAGAGCAGCAGCCCATCTGAGAGACTGTGTATAGAGCAGAGGAGCGGACAGACGGCAGC	5521
Db	5461	GTGGAGCAGCAGCCATCTGGAGACTGTGTATAGAGCAGAGCGGACAGACGGCAGC	5520
QY	5522	AGCGCGGGGGGTGGGGGAGCAGCAGCCGCCCGCCTCCACTCCATGCCATGCCACAGCAC	5581
Db	5521	AGCGCGGGGGGTGGGGGAGCAGCAGCCGCCCGCCTCCCACTCCATGCCACAGCAC	5580
QY	5582	TGCGCCATCTCCCTCTCGGAGCCAGGATGCTCTTCAGAGAACCGAGTGTGTTCACAA	5641
Db	5581	TGCGCCATCTCCCTCTCGGAGCCAGGATGCTCTTCAGAGAACCGAGTGTGTTCACAA	5640
QY	5642	ACAGGCATGAGAGGATCATCAACGCTGTGAGAGCCAGAGGCCACGGCTCTGAGGTCC	5701
Db	5641	ACAGGCATGAGAGGATCATCAACGCTGTGAGAGCCAGAGGCCACGGCTCTGAGGTCC	5700
QY	5702	ACCTTCACCTCTCTCAACCGGTTGCGCCAGCTGCACATTTCCACTGTGCACCCACTGCCA	5761
Db	5701	ACCTTCACCTCTCTCAACCGGTTGCGCCAGCTGCACATTTCCACTGTGCACCCACTGCCA	5760
QY	5762	CTGGCGCGGACCTCTGATAGGGGTCTAACCTTACCTCATGAGAGCGGCTTGTGCGCCAA	5821
Db	5761	CTGGCGCGGACCTCTGATAGGGGTCTAACCTTACCTCATGAGAGCGGCTTGTGCGCCAA	5820
QY	5822	GAGGCCCCCGGGTGCCTCGCGCAGAGCGGCCCCGAGCAGACACCGGCAATGCTTCTTC	5881
Db	5821	GAGGCCCCCGGGTGCCTCGCGCAGAGCGGCCCCGAGCAGACACCGGCAATGCTTCTTC	5880
QY	5882	GCCAAAGCCCCCAGCCCCGCTCGGGGTGTAGAGCCCCGCTCTCTCCCCGAGAAAGGCTGGAG	5941
Db	5881	GCCAAAGCCCCCAGCCCCGCTCGGGGTGTAGAGCCCCGCTCTCTCCCCGAGAAAGGCTGGAG	5940
QY	5942	CCCCGGCCCCAGTGTCTCTGTCTGTGACAGCGCACCATGCGCCGCAACCCCTGCGAAG	6001
Db	5941	CCCCGGCCCCAGTGTCTGTCTGTGTGACAGCGCACCATGCGCCGCAACCCCTGCGAAG	6000
QY	6002	AACTCTGCACCTCACCAAGCCAGCCCGGACCCGCGGGCGCACCTGCTGTGCTTGGAC	6061
Db	6001	AACTCTGCACCTCACCAAGCCAGCCCGGACCCGCGGGCGCACCTGCTGTGCTTGGAC	6060
QY	6062	CGGACCGGGGAAAAAGCATCAAAAGTAAACCTTTTCCATCCAGGAACTGGAACTCCGTTCT	6121
Db	6061	CGGACCGGGGAAAAAGCATCAAAAGTAAACCTTTTCCATCCAGGAACTGGAACTCCGTTCT	6120
QY	6122	CTGGATTACACGCGCAGAGCTTACGCCCCCGAAGGGGTGAAGCCGCTGAGCCCTGTGAGC	6181
Db	6121	CTGGATTTTACACGCGGTTTTTACGCCCCCGAAGGGGTGAAGCCGCTGAGCCCTGTGAGC	6180
QY	6182	TCACCCACTCTGACCCACGACAGAGGGGCTCCCCAAGACCTTGAAAGACTTCGACAAAGAC	6241
Db	6181	NN	6240
QY	6242	CACCTGAGGGGAGCTCGCGCCAAAGCAGAGCCCGCTGAAGCTTGGCGGGAGGCC	6301
Db	6241	NN	6300
QY	6302	GCCCACTTCCACACTGTGGGCGCTGCTGTGAAGCAGAGCCCTGTCTCAAGCCGTGTCTC	6361
Db	6301	GCCCACTTCCACACTGTGGGCGCTGCTGTGAAGCAGAGCCCTGTCTCAAGCCGTGTCTC	6360
QY	6362	CAGACCGGCGGAGGGGTCAAAGGTCAACAGCGGGGTGTCACCTGGCCAGCAGCATCAGT	6421
Db	6361	CAGACCGGCGGAGGGGTCAAAGGTCAACAGCGGGGTGTCACCTGGCCAGCAGCATCAGT	6420
QY	6422	GAGGTCAATCAACAGAGACTACCGCGCACCAACCAAGCAGACTGACGCAACCTGTGCC	6481
Db	6421	NN	6480
QY	6482	GCCCCCTCTACTCTTCCCTGTGGGCGCAGCTGCCCGTCTTGAACTCCGCGGCCACCC	6541
Db	6481	NN	6540
QY	6542	AGTGAAGCTCTACTCCGCGCCCGGAGCAATGTGTGCCCGGCGGCTTCCCCCAGC	6601



[illegible]

TITLE	JOURNAL	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	ORIGIN
Adams,M.D. and Cargill,M.	Interfering nonneutral evolution from human-chimp-mouse orthologous	gene tribes	science 302 (5652), 1960-1963 (2003)	14671302	2 (bases 1 to 6836)	Clark,A.G., Ghanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,A.G., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.	Direct Submission	Submitted (16-NOV-2003) Celera Genomics, 45 west Gude Drive, Rockville, MD 20850, USA	This sequence was made by sequencing genomic exons and ordering them based on alignment.
Location/Qualifiers	1..6836	/organism="Pan troglodytes"	/mol_type="genomic DNA"	/db_xref="taxon:9598"	<1..-56836	/gene="NCOR2"	/locus_tag="HCM4640"		
Query Match	49.2%	Score 4216.6;	DB 10;	Length 6836;					
Best Local Similarity	62.1%	Pred. No. 0;							
Matches 4243;	Conservative	0;	Mismatches 2593;	Indels	0;	Gaps	0;		
Db	79	CCGAGTCAGATGCGCCCGGAGCGACACAGGAGTGGGGCTCTGTGAGTACAGACACTC	138						
1	CCGAGTCAGATTGCCCGGAGCGACACAGAGTGGGGCTCTGTGAGTACAGACACTC	60							
Qy	139	CCGCGACTTATGGCTCTCCCACTGTGTGGCGGGCTCATCATCCAGGCCCGCGGAGGCC	198						
Db	61	CCGCGACTTATGGCTCTCCCACTGTGTGGCGGGCTCATCATCCAGGCCCGCGGAGGCC	120						
Qy	199	CTCCCTCTGTGTGAGTTCGACGCCGGGAAATGAACGGTCCAGAGAGCTCCACTGCGGCC	258						
Db	121	CTCCCTCTGTGTGAGTTCGACGCCGGGAAATGAACGNNNNNNNNNNNNNNNNNNNNNN	180						
Qy	259	AGAGTCCACTCATATCCTGCGCCGAGCTGGGGAATCAGAGATGAGTTCAATTGAAGCA	318						
Db	181	NN	240						
Qy	319	GCGCCCTCGGCTAGAGCTGCTGCTGACCCCTGTCTGCAACCTGCTGCGCAC	378						
Db	241	NN	300						
Qy	379	GGGCGAGCTGCGGGATCTGAAGACTCACCAAGACCGTAGCCTGAACGGCAAGCTGA	438						
Db	301	NN	360						
Qy	439	ACCGGTGTCTCCGCCAGCCCCCGGACACTGACCTGAGCTGAGCTGTGCGGCCACG	498						
Db	361	ACCGGTGTCTCCGCCAGCCCCCGGACACTGACCTGAGCTGAGCTGTGCGGCCACG	420						
Qy	499	GCTGTCCAAGAAGAAGTGTATCCAGAACATGAGACCGGTGGAACGAGAGATCAACATGT	558						
Db	421	GCTGTCCAAGAAGAAGTGTATCCAGAACATGAGACCGGTGGAACGAGAGATCAACATGT	480						
Qy	559	AGAGCAGCAATCTCTTAAGCTGGAAGAAGACAGCAACAGCTGAGAGAGAGCTGCCA	618						
Db	481	AGAGCAGCAATCTCTTAAGCTGGAAGAAGACAGNNNNNNNTTGAAGAGAGAGGCTNCCA	540						
Qy	619	GCGGCCGAGCTTGAGAAACCGGTGTCAACGCGGCCCATGTGAGTTCGAAGCACCGAGCTT	678						
Db	541	GCGGCCGAGCTTGAGAAACCGGTGTCAACGCGGCCCATGTGAGTTCGAAGCACCGAGCTT	600						
Qy	679	GGTGCAGATCATCTAAGCAGAGAAACGGGAAGAGGCTGAAGCTCAGCATCGGATTCGGA	738						
Db	601	GGTGCAGATCATCTAAGCAGAGAAACCGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	660						

[illegible]

Db	1741	CATGAGACTGATAGAGTTCTTCCTGCAAGAAAGAAATGAAACAGCCAAAGAAAGG	1800
Qy	1879	TCCTCTGGAACACGCGCCGCAACTGTCGGCCATGCGCCGATGCTGAGCTCAAGACTGT	1938
Db	1801	TCCTCTGGAACACGCGCCGCAACTGTCGGCCATGCGCCGATGCTGAGCTCAAGACTGT	1866
Qy	1939	GTGCGACTGTAAAGACTTCTACTTCAACTACAGAGAGGCAAACTCGATGAGACTTT	1998
Db	1861	GTCCCACTGTAAAGACTTCTACTTCAACTACAGAGAGGCAAACTCGATGAGACTTT	1920
Qy	1999	GCACACACACAAAGCTGAAGATGAGAGAGAGAAACGCGCGAGAGAAAGAAAGAAAGC	2058
Db	1921	GCACACACACAAAGCTGAAGATGAGAGAGAGAAACGCGAGAGAGAAAGAAAGAAAGC	1980
Qy	2059	GCCGCGCGCGGCGCAAGCAGAGAGAGCTGCATTCCGCCGTGTGAGAGATGAGAGATGGA	2118
Db	1981	GCCGCGCGCGGCGCAAGCAGAGAGAGCTGCATTCCGCCGTGTGAGAGATGAGAGATGGA	2040
Qy	2119	GGCGTCTGGGCGTGAGCGGAAATGAGAGAGAGATGCTGAGAGAGCTGAAGCTTACATGC	2178
Db	2041	GGCGTCTGGGCGTGAGCGGAAATGAGAGAGAGATGCTGAGAGAGCTGAAGGNNNNNNNNNN	2100
Qy	2179	CTCTGAGAAATGAGGTGCCCAGAGGGGAAATGAGTGGCCAGCCACTGTCAACAACAGCTC	2238
Db	2101	NN	2160
Qy	2239	AGACACCGAGAGACTCCCTCTCTCTCACTGAGGCGCCAGAGACACAGGCGAGAAATGG	2298
Db	2161	NN	2220
Qy	2299	GCCCCAAGCCCCCAGACCACTCGGGCGCCGACGGGCAACCCCAGAGCCACCCACCCAC	2358
Db	2221	NN	2280
Qy	2359	ACGGAAGACATCCCGGGCCCCATTGAGCCACCCCGGCTCTGAAGCACCCGAGCCCC	2418
Db	2281	NN	2340
Qy	2419	TAGCCCCCACAAGACCCCCATGCGCTCTGCACTCTCTGTGTGCCCAAGAGAGA	2478
Db	2341	NN	2400
Qy	2479	GAAAGAGAGAGAGACGACAGACGCCCCCATGAGAGAGGAGAGAGACGAAAGCCCC	2538
Db	2401	NNAGAGAGAGAGAGACGACAGACGCCCCCATGAGAGAGGAGAGAGAGAGAGAGCCCC	2460
Qy	2539	CGCGGCTGAGAGAGCTGGCAGTGGACACAGGGAAGGCCGCTCAAGAGCCAGTGT	2598
Db	2461	CGCCNNNAGAGAGCTGGCAGTGGACACAGGGAAGGCCGCTCAAGAGCCAGTGT	2520
Qy	2599	CACGAGAGAAAGCGAGAGAGGGGCGGCGCAAGGGCAAGAGCGCGAGAGCGCTGAGGCGAC	2658
Db	2521	CACGAGAGAAAGCGAGAGAGGGGCGGCGCAAGGGCAAGAGCGCGAGAGCGCTGAGGCTAC	2580
Qy	2659	GCGCGAGAGAGGCGCTCAAGGACGAGAAAGAGAGAGGCGAGAGCGGCAAGGCGCAACTGC	2718
Db	2581	GCGCGAGAGAGGCGCTCAAGGACGAGAAAGAGAGAGGCGAGAGCGGCAAGGCGCAACTGC	2640
Qy	2719	CAAGAGCTCGAGCGCGCCCCCAAGACAGCAGACTCCAGTGTCTACTCTGACAGTCAAGAGT	2778
Db	2641	CAAGAGCTCGAGCGCGCCCCCAAGACAGCAGACTCCAGTGTCTACTCTGACAGTCAAGAGT	2700
Qy	2779	GGAAGAGAGCGGAGGCGGCGGCAAGAAACCGGCTGTCTCCCAAGGCGCAAGCTCTCTAC	2838
Db	2701	GGAAGAGAGCGGAGGCGGCGGCAAGAAACCGGCTGTCTCTCCCAAGGCGCAAGCTCTCTAC	2760
Qy	2839	CCCGACTGGGCGACCCCGGGGCAATGTGCTCACTCCAGAGAGCACTGGACTGGAAGCACT	2898
Db	2761	CCCGACTGGGCGACCCCGGGGCAATGTGCTCACTCCAGAGAGCACTGGACTGGAAGCACT	2820
Qy	2899	GAAGCAGCAGAGCGGCTGCATCCCCCCATCTCAGGTCAACAAATGCAAGAGCCCCCGG	2958

Db 2821 GAAGACGAGCGGCTGCAATCCCCCATTCAGGTCAACCAAGTCCATGAGCCCCCG 2880  
QY 2959 GGAGAGCGAGCTCCACCAAGCAGTCCCCAGGCCCAACGCGCAACCGCAAAACCTGCA 3018  
Db 2881 GGAGAGCGAGCTCCACCAAGCAGTCCCCAGGCCCAACGCGCAACCGCAAAACCTGCA 2940  
QY 3019 GCGGAGAGCGAGCGCCCTCAGAGCCTGAGCAGAGCCCCCGGGGAGAGAGAGCC 3078  
Db 2941 GCGGAGAGAGAGCGCCCTCAGAGCCTGAGCAGAGCCCCCGGGGAGAGAGAGCC 3000  
QY 3079 GGCACCCCCGCGCAAGAGAGCCTTCGACCGCAGAGCCAGAGAGCTGCTGGGAGCC 3138  
Db 3001 GGCACCCCCGCGCAAGAGAGCCTTCGACCGCAGAGCCAGAGAGCTGCTGGGAGCC 3060  
QY 3139 CCTTGCTGGACTTCGGGCTGCTTCCTCCGTCGCCCTCGAGGTGTCAGAGCCTC 3198  
Db 3061 NNN 3120  
QY 3199 CCGGCAAGCGCCGAGACCCCTCAGCCTTCTCTACGCTCCACCTGTCAACCACTGCCCT 3258  
Db 3121 NNN 3180  
QY 3259 GGGCTTCATGACATGCGCCCGGCGCTCTGCGCGGCCCAACCAATCTTCAACCGCC 3318  
Db 3181 GGGCTTCATGACATGCGCCCGGCGCTCTGCGCGGCCCAACCAATCTTCAACCGCC 3240  
QY 3319 TCCCTCATCTCTCTGCGCAAGACCGCGCTCTGAGGCAATAGTGGCACTCTC 3378  
Db 3241 TCCCTCATCTCTCTGCGCAAGACCGCGCTCTGAGGCAATAGTGGCACTCTC 3300  
QY 3379 CCAAGGAATGTCGATCAGCTCCAGTCCCGTACTAGAGCATGCCAAGGCCCGGTGAG 3438  
Db 3301 CCAAGGAATGTCGATCAGCTCCAGTCCCGTACTAGAGCATGCCAAGGCCCGGTGAG 3360  
QY 3439 CCCTGTACCATAGGAGCTGCGCTGCGCATGAGACCCCAAAAGCTGCAACCTTCAAGCG 3498  
Db 3361 CCGCTGTACCATAGGAGCTGCGCTGCGCATGAGACCCCAAAAGCTGCAACCTTNNAGCG 3420  
QY 3499 AGTGAAGCAGAGCAGCTGTCCCAAGGGGCGAGGCTGGGGCCACCGGAAGAGCTGGGGGT 3558  
Db 3421 AGTGAAGCAGAGCAGCTGTCCCAAGGGGCGAGGCTGGGGCCACCGGAAGAGCTGGGGGT 3480  
QY 3559 GCCCAGGCCAGAGAGCGTCCGTGCTGAGAGGAGCAGCTGTGGGTCAAGTCCGGAGCG 3618  
Db 3481 GCCCAGGCCAGAGAGCGTCCGTGCTGAGAGGAGCAGCTGTGGGTCAAGTCCGGAGCG 3540  
QY 3619 AAGCATCACCAAGGCAATTCAGCAACAGGAGTCCCTCGGACAGCGCCATCATACCG 3678  
Db 3541 AAGCATCACCAAGGCAATTCAGCAACAGGAGTCCCTCGGACAGCGCCATCATACCG 3600  
QY 3679 CGGCTTCATCAACCAAGGCAATTCAGCAACAGGAGTCCCTCGGACAGCGCCATCATACCG 3738  
Db 3601 CGGCTTCATCAACCAAGGCAATTCAGCAACAGGAGTCCCTCGGACAGCGCCATCATACCG 3660  
QY 3739 CATCGGAGAGCAGCGCGAGTGGTGGACGCGGCGGAGAGCAGCGTGGCCAAAGG 3798  
Db 3661 CATCGGAGAGCAGCGCGAGTGGTGGACGCGGCGGAGAGCAGCGTGGCCAAAGG 3720  
QY 3799 CCAAGTCACTAGAGGCAAGAGGCAACGCTTGTCTATGAGGAGTGTCTGT 3858  
Db 3721 CCAAGTCACTAGAGGCAAGAGGCAACGCTTGTCTATGAGGAGTGTCTGT 3780  
QY 3859 GACCCAGTCTCAAGAGAGAGCGGCAAGAGCAGTCAAGACCCCCCATGAGAGCGCCG 3918  
Db 3781 NNN 3840  
QY 3919 CCCCAAGCGCACTATGATGAGAGGCGCGTGGGAGAGCAGCTCTCTCAGCAG 3978  
Db 3841 NNN 3900  
QY 3979 CATCGAAGTCTCATAGGCGCTGTCATCCCGCGGAGCGACAGCGCCCAACCTCAA 4038  
Db 3901 NNN 3960

QY 4039 AGAGCAGACCAATCCGGGGTTCATCAACAAGGATCCCTCGTCTACGTGAGGC 4098  
Db 3961 AGAGCAGACCAATCCGGGGTTCATCAACAAGGATCCCTCGTCTACGTGAGGC 4020  
QY 4099 ACAGAGAGCTACTGCTCGGAGGCGCAAGCTCTTAAAGCGGAGGCGAGCTTCGCC 4158  
Db 4021 NNN 4080  
QY 4159 CCAACCGCTCAAGGAGCTGACGAGGCTTCAAGAGCAGAGGCCCTGGGCCCTGA 4218  
Db 4081 NNN 4140  
QY 4219 GCTGAAGCGGCGCATGAGGCGCTGTGCGCAAGGTGAAGAGGCGCGCTCATCA 4278  
Db 4141 NNN 4200  
QY 4279 TGAATCCGCGCGAGAGCTGCGGCAACGCCGAGCTGCCCTGGCCCGCGCT 4338  
Db 4201 NNN 4260  
QY 4339 CAAAGAGGCTCATCAAGCAGGACCCCGCTCAAGTACAGACCGGCGCTCAAC 4398  
Db 4261 NNN 4320  
QY 4399 TGGCTCAAAAGCAGAGCTAGCTCCCTCATCGGAGCGCGGCGGAGCTTCCAC 4458  
Db 4321 NNN 4380  
QY 4459 CGTGACCCGCTGATGATGAGCGAGCGCGGCGCATGGAAGTGTCTTCAAGAG 4518  
Db 4381 CGTGACCCGCTGATGATGAGCGAGCGCGGCGCATGGAAGTGTCTTCAAGAG 4440  
QY 4519 GAGCTGAAGAGCGGCGAGGAGCGCGCAGAGCTTGGGGGGTCCATTGGCGCGCC 4578  
Db 4441 GAGCTGAAGAGCGGCGAGGAGCGCGCAGAGCTTGGGGGGTCCATTGGCGCGCC 4500  
QY 4579 CCGGCTATTGCTGAGCTGAGGTAAAGCGGCGGAGAGCGCCCTGACCTATGAGAGCA 4638  
Db 4501 CCGGCTATTGCTGAGCTGAGGTAAAGCGGCGGAGAGCGCCCTGACCTATGAGAGCA 4560  
QY 4639 CGGGGCAACCTTTCGCGGCACTCCCAAGAGTTCCCGTGAACATGCGGAGCGCA 4698  
Db 4561 CGGGGCAACCTTTCGCGGCACTCCCGGAGTTCCCGTGAACATGCGGAGCGCA 4620  
QY 4699 GCGGCGCTGCAAGAGGAGGCTTTCGTCAAGCAAGGATCCAGAGACCGAAAGCTGAC 4758  
Db 4621 GCGGCGCTGCAAGAGGAGGCTTTCGTCAAGCAAGGATCCAGAGACCGAAAGCTGAC 4680  
QY 4759 GTGAGCGCTGTGAGATGCGCAAGTCCCGCAGACAGCTGCGCGAGACCAACCA 4818  
Db 4681 NNN 4740  
QY 4819 CCCCATCTCGCCTATGAGCACTGCTTGGGGCGTGAAGTGAAGCTGTATGAGAG 4878  
Db 4741 NNN 4800  
QY 4879 CCAATCCCTGCGCTTGAACCCCACTCTAACCAGCGGATCTCTGTGAAGCAGC 4938  
Db 4801 NNN 4860  
QY 4939 CGCTGCTTACTGCTGCGCCGAGCACTGCGCCGCAACCCCACTACCGGAGCTGTAGCC 4998  
Db 4861 CGCTGCTTACTGCTGCGCCGAGCACTGCGCCGCAACCCCACTACCGGAGCTGTAGCC 4920  
QY 4999 ACCCTACCTCATCGCGGCTTACCCGAGCAGCGGCGTGAAGAACCGGAGACCATCAT 5058  
Db 4921 ACCCTACCTCATCGCGGCTTACCCGAGCAGCGGCGTGAAGAACCGGAGACCATCAT 4980  
QY 5059 CATATCTATCATCTGCGAGAGATGAGCAACAAGGCAACCGCATGCGCCAGG 5118  
Db 4981 CATATCTATCATCTGCGAGAGATGAGCAACAAGGCAACCGCATGCGCCAGG 5040

Oy	5119	AGCTGATATGCTGAGGGGCTCTGCGCCCGGAGTCTCGTGGACATCACTACGCTGC	5178
Db	5041	AGCTGAATGCTGAGGGGCTCTGCGCCCGGAGTCTCTGCTGGACATCACTACGCTGC	5100
Oy	5179	GGGTCCCGGAGGACATCATGCACTTGTCCCAAGTGCACACTGCCTGTGCTCGATCCCC	5238
Db	5101	GGGTCCCGAGGNN	5166
Oy	5239	GACACACAGGACCCCGAGCAACGCAATGAAACGGCTTGCTACTCCCAACGGGCCCA	5298
Db	5161	NN	5220
Oy	5299	GCCCTTGACAGCGCGCACAGAGCTCCCACTCTCCCAAGAGTCCAAACACTTGAC	5358
Db	5221	NN	5280
Oy	5359	AAACCAACACACACAGTCTCGTCCGACCGGAGGAGACCGGAGTCGAGCGGACCG	5418
Db	5281	MAACCAACACACACATCTCTGTCGACCGGAGGAGACCGGAGTCAGAGCGNNNNNN	5340
Oy	5419	GGATCGGAGCGGAAAATCCTCATCTCAAGTCCACACGAGTGAAGACACGACCCAT	5478
Db	5341	NNNNNNNAGCGGGAAGAACCATCTCTCACGTCACANNNNNNNGTGAAGCAACGCCCAT	5400
Oy	5479	CTGGAAGACTGTGTACAGACAGAGAGGGGACAGCGCGACAGCGCGCGGGGGTGGGG	5538
Db	5401	CTTGAGNNNNNGTACAGACAGAGAGCGGACAGCGACAGCGGTGGGGGTGGGGG	5460
Oy	5539	CAGCAGACCGCGCCCGCTCCCATCTCCCATAGCCCAACAGACACTCGCCCTCGC	5598
Db	5461	CAGCAGACCGCGCCCGCTCCCATCTCCCATAGCCCAACAGACACTCGCCCTCGC	5520
Oy	5599	GACCCAGATGCCCTCAGACAGACCCAGTGTCTTCAACACAGGACATGAAGGGTAT	5658
Db	5521	NN	5580
Oy	5659	CATCACCGCTGTGAGGCCAGACAGCCACGGTCTTGAGTTCACCTCACCTCTCAC	5718
Db	5581	NN	5640
Oy	5719	CGTTCGCCACGTGCACATTCACACTGCAACCTGCCACTGGGGGGCAACCTCGA	5778
Db	5641	NN	5700
Oy	5779	TGGGGTCTAACCTTACCCTCATGAGCCGCTCTTGCTGCCAAGAGAGCCCCGGGTGC	5838
Db	5701	NN	5760
Oy	5839	CCGGCCAGAGCGGCCCCGAGACACACGGGCAATGCTTCTCGCAAGCCCAACGCC	5898
Db	5761	NN	5820
Oy	5899	CTCCGGGTGAGGCCCGCTCTCCCCAGCAAGGCTCGGAGCCCGGCCCCCTATGTCC	5958
Db	5821	NN	5880
Oy	5959	TCTGTCTCTGGACAGCACATTCGACCGCACCCCTCGAAGAACTCGACCTCACCA	6018
Db	5881	CCCCGTCTTGCCCAAGCACATGCGCCGACCCCTCGAAGAACTCTCGACCTCACCA	5940
Oy	6019	CGCCAGCCCGACCCCGCGGCGCACCTGCTCGGCTCGACCCCGCACCGGAAAAAGAC	6078
Db	5941	CGCCAGCCCGGACCCCGCGGCGCACCTGCTCGGCTCGACCCCGCACCGGAAAAAGAC	6000
Oy	6079	TCAAAGTAAACCTTTTCATTCAGAACTGGAACCTCGTTCTTGAGTTACACGCGAG	6138
Db	6001	TCAAAGTAAACCTTTTCATTCAGAACTGGAACCTCGTTCTTGAGNNNNNNNNNNNN	6060
Oy	6139	CAGCTACAGCCCGGAAGGGGTGAGCCGCTGAGCTGTGAGTCAACCACTGACCCA	6198
Db	6061	NN	6120
Oy	6199	CGACAGGGGCTCCCAAGCACTTGAAAGACTGCAACAAGACACTGAGGGGGAGCT	6258

[illegible]

Ferriere, S., Wang, G., Zheng, X.H., White, T.J., Sinsky, J.J.,  
Adams, M.D. and Cargill, M.  
Direct Submission  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
COMMENT Location/Qualifiers  
FEATURES  
source 1..7013  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
<1..>7013  
/gene="NCOR2"  
/locus\_tag="HMC4640"

## ORIGIN

Query Match 38.7%; Score 3315.2; DB 10; Length 7013;

Best Local Similarity 58.3%; Pred. No. 0;

Matches 4299; Conservative 0; Mismatches 2711; Indels 365; Gaps 14;

```
QY 2 ATGTGGGCTCCACACAGCTTTGTGACAGAGTGAAGGSCACTGAGCCCGCTACCG 61
DB 1 ATGTGAGATCCACACAGCTGTGGACAGACATGGGGGCTGTGAGCCCGCTACCG 60
QY 62 CCCACAGCCTTCTACCCAGTGAATGGCCGAGCAGCAGGAGCGTGGGCTCTG 121
DB 61 CCCATGAGATCTCTCAATCCGATGAGCCGGTCCACAGGAGTGGGGCTGCTT 120
QY 122 GAGTACAGACACCACTCCGCGACTATGCTCTCCACTGTGGCGGGCTTCATCTCA 181
DB 121 GAGTACCAACACCAACCCCGTGACTACCTCAACCTGTCAACCCGGTTCATCTCA 180
QY 182 CCCACAGGCGAGGCGCTCTCTGTGTCTGAGTTCAAGCCCGGAATGAAGCTCCG 241
DB 181 CCACAGAGAGGGGCGCTCTCACTGTGTGAGTTCAAGCTGGAGTGAAC-GNNNNN 239
QY 242 GAGCTCACCTGGGCGAGAGTCCACTACTCTGCGGAGCTGGGGAAGTACAGATG 301
DB 240 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 299
QY 302 GAGTTCAATTGAAGCAAGCGCCCTCGGCTAGAGCTGCTGACCCCTGTGCGACG 361
DB 300 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 359
QY 362 TCAACCTCTGTGCGCACGCGGACAGCTGGGGAATCTGAAGACTCAACAGACCT 421
DB 360 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 419
QY 422 CTGACGGGGAAGCTGGAACCGGTGTCTCCCGGAGCCCGGCACTGACCTGAGCT 481
DB 420 CTGGACGGGAAGCTGGAAGCTGTGTCACTCCAGTCCCGGACGCTGACCTGAGCT 479
QY 482 GAGCTGTGCTCCGCAAGCTGTCTCAAGAGAGAGCTGATCCAGAACTGACCGCTG 541
DB 480 GAGCTGTGCGCACTGTGACTGTCTCAAGAGAGAGCTGATCCAGAACTGACCGCT 539
QY 542 CGAGAGATCACTGTTAGAGAGAGAGTCTTAAGCTGAAGAAAGACAGCAAGCTG 601
DB 540 CCGTGAATCACTGTGTAGAGAGAGAGTCTCCAACTGAAGAAAGAGAGCAAGT 599
QY 602 GAGGAGGAGCTGCGCAAGCGCCGAGCTGAGAAACCGGTGTCAACCGCGCCAT 661
DB 600 GAGGAGGAGCGCGCAAGCGCCGAGACCGGAGAAACCGGTGTGCGCAACCGCAT 659
QY 662 TCGAAGACACGAGCTGTGTGAGATCATCTACAGAGAGAACCGGAAAGAGCTGA 721
DB 660 TCGAAGACACGAGGCTGTGTGAGATCATCTACAGAGAGAACCGGNNNNNNNN 719
QY 722 GCACATCGATTCTGGAAGGCTGGGGCCCAAGTGAAGCTGCGCTGTACACAGCC 781
DB 720 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 779
```

```
QY 782 TCCGACACCCGCGAGTATCATGAAATCAATAATAACAGGCGATGCGGAAGACTA 841
DB 780 TCTGACACACGCGAGTACATGAAACATGAAATTAACAGGCGATGCGGAAGACTG 839
QY 842 ATCTTGACTTCAAGAGAGAAATCACGCTCGAATAATGAAAGCAAGTTCTGCG 901
DB 840 ATCTTGACTTAAAGGAGAAACACGCGCGAAGAGTGGAAACAGCCCTTCTGCG 899
QY 902 CGCTATGACAGCTCATGAGAGGCTTGGAAAAAAGTGAAGGACATCGAAAAAACC 961
DB 900 CGCTATGACAGCTCATGAGAGGCTTGGAAAAAAGTGAAGGACATGAGAAACAT 959
QY 962 CGCGCGCGGCGCAAGAGAGCAAGTGTGCGAGTACTACGAAAGCATTCCTGAGAT 1021
DB 960 CGAAGAGAGGCGCAAGAGAGCAAGTGAAGGATTACTACGAGAAACAGTTCCG 1019
QY 1022 CGCAAGACGCGAGCTGCGAGAGCGCATGACAGAGAGTGGGCCAAGCGGCGAG 1081
DB 1020 CGCAAGACGCGAGCTGCGAGAGCGCATGACAGAGAGTGGGCCAAGCGTGGC 1079
QY 1082 CTGTCAATGTGGGCGCGCGCGAGCGAGCAGAGGTGTCAAGATGCAATGAGCT 1141
DB 1080 CTGTCAATGTGGGCGCGCGCGAGCGAGCAGAGGTGTGTGAATGCAATGAGCT 1139
QY 1142 GAGCAGAGAACTTGAAGAAAGCATGCGCAGCTGCGCTGATCCCGCATGCTGAC 1201
DB 1140 GAGCAGAGAACTTGAAGAAAGCATGCGCAGCTGCGCTGATCCCGCATGCTGAC 1199
QY 1202 GAGCTGACCAAGAGCGCATCAAGTTTCAATCAATGAACGGGCTTATGCGGAC 1261
DB 1200 GAGCGGAGCAAGAGAGATCAAGTTTCAATCAATGAATGAGTCAATGATGACCC 1259
QY 1262 AAGGTGAACAAAGCCGCAAGTCAATGAATGAGAGTGAAGAGAGAGAGCTTC 1321
DB 1260 AAGGTGAACAAAGCGGTGATCAATGATGAGAGTGAAGAGAGAGAGAGAGCT 1319
QY 1322 CGGAGAAAGTTCAATGAGATCCCAAGAACTTGGCTGATCGCATCATTCCTG 1381
DB 1320 CGGAGAAAGTTCAATGAGATCCCAAGAACTTGGCTGATCGCATCATTCCTG 1379
QY 1382 AAGCAGTGGCTGAGTGGCTCTTATTAATTAATTAATTAATTAATTAATTAAT 1441
DB 1380 NNNNACGTCGTCGATGATGTCCTTATTAATTAATTAATTAATTAATTAATTA 1439
QY 1442 AGCTGTGAGACGAGCTATGCGCGCGGCAAGAGCAGCAGCAACAGCAGCAG 1501
DB 1440 AGCTGTGAGAGGAGCTATGCGCGCGGCAAGAGCAGCAGCAGCAGCAGCAG 1499
QY 1502 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1561
DB 1500 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1559
QY 1562 GATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621
DB 1560 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1619
QY 1622 GACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681
DB 1620 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1679
QY 1682 AAGGAGGCTGTGGCTCAAGAGCGCAAACTGCAAGCAGCAGGAGAGAGAGAG 1741
DB 1680 AAGGAGGCTGTGGCTCAAGAGCGCAAACTGCAAGCAGCAGGAGAGAGAGAG 1739
QY 1742 CGCATCACCGCTCAATGAGCTAATGAGCCAAAGGAGAGAGAGAGAGAGAG 1801
DB 1740 CGCATCACCGCTCAATGAGCTAATGAGCCAAAGGAGAGAGAGAGAGAGAGAG 1799
QY 1802 AGCGCGAGCTGCGCTCAATGAGCTAATGAGCTAATGAGCTAATGAGCTAAT 1861
DB 1800 AGTTCAAGAGCTGCTCAATGAGCTAATGAGCTAATGAGCTAATGAGCTAAT 1859
QY 1862 GAAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1921
```

Db	1660	GAGACAGCAAAAGAAAGGCGCTCTCGAATGAGAGAACTGTGACGCCATTGCGCAGTGG	1919
Qy	1922	GTGGAGCTCCAAAGACTGTGCGCAGTGTAAAGACTTCTCACTTCAACTACAAAGAGGAG	1981
Db	1920	GTGGAGCTCCAAAGACCGTGTCCCAAGTGTAAAGACTTCTCACTTCAACTACAAAGAGGAG	1979
Qy	1982	AACTCGATGAGATCTTGACAGCAGCAAGCTGAAGATGAGAAAGAGAGAAAGCGCGCG	2041
Db	1980	AACTCGACGAATCTTTCAGCAGCAAGCTGAAGATGAGAAAGAGAGAAAGCGCTCGG	2039
Qy	2042	AGGAAGAAAGAAAGAGCGCCCGCGCGGCCAGCCAGAGAGCTGCATTCCCGCGTGTG	2101
Db	2040	AGGAAGAAAGAAAGAGCGCCCGCGCGGCCAGCCAGAGAGCTGCATTCCCGCGTGTG	2099
Qy	2102	GAGATGAGAGATGAGAGCGCTCGGCGCTGAGCGGAAATGAGAGAGATGATGTGAGAG	2161
Db	2100	GAGACGAAAGATGAGAGCATCAGCGCCCAAGTGCATGAGAGAGCTGCGGAGAG	2159
Qy	2162	GCTGAGCTTATACATGCTCTTGGAATGAGGTGCCAGAGGGGAATGCATGTGCCACGC	2221
Db	2160	GCAAGAGNNCT	2219
Qy	2222	ACTGTCAACAACAGCTCAGACACCGAGAGCATCCCTCTCTCTCACTGAAGCGCCAG	2281
Db	2220	GCT-----CGAGCTCCCAAGAGATGCCAAG	2249
Qy	2282	GACACAGGAGAGATGGGCGCAAGCGCCCAAGCCTCGGCGCGAGCGGAGCAACCCCA	2341
Db	2250	GAGC-----	2253
Qy	2342	GCGCCACCCACCCACCGAGAGCATCCGCGCGCCCATTTAGGCCACCCCGGCTCT	2401
Db	2254	-----	2253
Qy	2402	GAACTCCGAGCGCCCTTACGCCCCCAACAGCACCCCATGCGCTCTGCACCTCTCT	2461
Db	2254	-----	2253
Qy	2462	GTGTGCCCAAGAGAGAGAGAGAGACCGCAGCAGCGCCCGCATGTGAGAGAGG	2521
Db	2254	-----AGAACTCTGAGCGCGAGAGATTCATGTGGAAAGCCAGAGA-----	2296
Qy	2522	GAGAGCAGAAAGCCCCCGCGCTGAGAGCTGCGCATGTGACAACGAGAAAGCCGAGAG	2581
Db	2297	-----GCCGAGGCTCTGAGAGACCCCGGAG	2324
Qy	2582	CCCTCAAGAGCGAGTGCACGAGAGACCGAGAGGGCC---GGCCAAAGGCAAGAC	2638
Db	2325	AGTGTAAAGATGACCAACAGAGAGAGACCGAGAGAGCTTGAAGCAAGGCCAAGGCG	2384
Qy	2639	GCGAGGCGCGCTGAGGCGCACCGCGCGAGGGGGCGCTCAAGGCAAGAAAGAGAGCGG	2699
Db	2385	ACAGAGGCGCATTTGAATCTGTCTGAGGACCATTTAAG---GTGAGAGAGCTGTATGC	2441
Qy	2699	AGCGCAGGGCACCATCTGCCAAGAGCTCGGAGCGCCCCCAAGCAGCGCATCCAGTGT	2758
Db	2442	AAAGCAGCTGTGACCAAGGCTTCCAGTCAAGGTGCACCCAGAGCATGTGATCCAGTGC	2501
Qy	2759	ACTTGCAGTGCAGACGAGTGTGATGAGGCGCGAGGGCGCGACAGAACCGGCTGTCTC	2818
Db	2502	ACTTGCAGTGCAGTGTGAGTGTGAGAACCCAGAGAGGTGACAAAGGCGAGGCTGTCTC	2561
Qy	2819	CCAAAGGCGCAAGCTCTCAACCCCGCATGAGGACCCCGGGCGCAATGCTTCAACCCAGAG	2878
Db	2562	CCAAAGGCGCAAGCTCTCAACCCCGCATGAGATCCCGGGCGCAAGTCTCTGCCCAAGAG	2621
Qy	2879	CCACTGTGACTGAAGCAGTGTGAAGCAGAGCGGCTGCATCCCGCATCAGGTCAAC	2938
Db	2622	CCGCTGTGACTGAAGCAGTGTGAAGCAGAGCGGCTGCATCCCGCATCAGGTCAAC	2681
Qy	2939	AAAGTCCATGAGCCCCCGGAGAGCGAGCTCCCAACCAAGCTTCCCGCAGCCCCA	2998

D	2682	AAGGTCATGAGCCCCCGGAGGAGACAGATACCCCAAGCAATTCCCTGTGCT	2741
Q	2999	CCGCGCACCGCAAAACCTGACGCGCGAGAGCGACGCCCTCAGACGCTGGCAGAGCCCC	3058
D	2742	CCACCCACGCGACGACCTTACAGCCAGAGGATGACGTGTCTCAGCAGTGGGAGGAAGTCCA	2801
Q	3059	CGGGGCAAGACGAGAGCCCGGCAACCCCGCGCAAGAGAGCCTTCGACCGGAGGCC	3118
D	2802	CGTGGCAAGTCCCGACGCGCAGTGCCTTCGCGCAGAAAGAGNNNNNNNNNNNNNNNN	2861
Q	3119	CAGAAAGCTGCGTGGGAGCCCCCTTGTGGACTTCGCGGCTGCCTTCCCGTCCCCC	3178
D	2862	NN	2921
Q	3179	CGTGAAGGTGATCAAGGCGCTCCCGCATGCCCCGAGCCCTCAGCCTTCTCTACGCTCA	3238
D	2922	NN	2981
Q	3339	CTGTGTACCCACCTGCCCCCTGGGCGCTCATATGACATGCCCCGCGCTCTGCGCGCCA	3298
D	2982	NNNNNGTACCCGCTGCGCTCTGGGCGCTCCACGATAGTCCGCGCCGCTCTGCCACG--T	3038
Q	3299	CCGACCATCTCCAAACCGGCTCCCTCATCTCCCTGCGCAAGACGCCAGGCTCTCGAG	3358
D	3039	CCCCCATCTTAAACCCCCACCCCTCATCTCTCTGCGACAGATCCCGGCGTACTTGAAG	3098
Q	3359	AGGCAAAATAGGTGCATCTCCCAAGAGATGTGCGTCCAGCTCCAGTCCCGTACTCAGAG	3418
D	3099	AGGCAAGCTGGGTGCATCTCCACAGGGATGTCAATCCAGCTTGGTGTGCTCACTCAGAG	3158
Q	3419	CATGCCAAGGCCCCGGTGGGCGCTGTCACTCAATGGGGCTGCCCTTGCCATATGACCCCAA	3478
D	3159	CATGCCAAGGCCCCCAATGGGCGCTCTCACTCATATGGGGCTGCCCTTGGCGTGGACCTTAAG	3218
Q	3479	AAGGTGCACCCCTTCAAGGAGTGAACAGGAGGAGCGTCCCAAGGGGCGAGGCTGGG	3538
D	3219	AAGCTGCACCCCTTCAAGTGAATGAAGAGAGAAAGTTGTCCCTCGGGGTCAAGGCTGGG	3278
Q	3539	CCACCGGAGAGCCTGGGGGCTGCCCCACAGCCGAGAGGCGTCCGTCTGAGAGGACAGCT	3598
D	3279	CCACTTGAAGTCTGGGAGTGGCCACTGCTCTCAGGAGACTCTGTGTGAGAGGAGACAGCA	3338
Q	3599	CTGGGCTCAGTTCCGGGCGGAAGCATCAACAAAGCATTTCCACAGACACGGGTGCGCTCG	3658
D	3339	CTGGGCTCCGCGCAACATGGAGATCAACCAAGGCGCTCCCGCACTACCCGGG-----CT	3392
Q	3659	GACAGCGCCATCAATATACCGCGGTTCATATACCCACGGGACGCCAGCTGATGCTGTAC	3718
D	3393	GCAGACGCGCCCAAGCTACAGAGGCTCTATATACCCACGGGACGCCGAGAGTCTCTTAC	3452
Q	3719	AAGGCAACATCAACAGATATCTGGGAGAGGACAGCCGAGTGCCTTGGACCGCGCGCG	3778
D	3453	AAGGCTACATCAGCAGAGATGTGTGGTGAAGGACGCCCAAGTCCCTTGAACGGGACGGA	3512
Q	3779	GAGGACAGCCTGCCCCAAGGCGCATGTATCTTACGAAGGCAAGAGGCGCATGTCTTGTCC	3838
D	3513	GAGGACACCTGCGCCCAAGGCGCATGTATCTTATAGGGCAAGAAAGGCGCATGTCTTATCC	3572
Q	3839	TATATAGGTGGCATGTGTCTGACCCCAAGTGTCCAGAGAGACGGCAGAAGAGCTCAGGA	3898
D	3573	TATATAAANN	3632
Q	3899	CCCCCCATGAGACGCGCGCCCCCAAGCGACCTATGACATGATGAGAGGCGCGGTGGGC	3958
D	3633	NN	3692
Q	3959	AGAGCCATCTCTCAGCAGCATGAAGTCTCATYGGCGCGTGCATCTCCGCGGAGCGA	4018
D	3693	NN	3750
Q	4019	CACAGCCCCCACCACTTCAAGAGACGACCACTCCGCGGGTCTATCAACAAGGATC	4078
D	3751	----ACAGCCCCCACTTCAAGAGACGACATCACTCCGAGGCTCTCATCAACCAAG-----	3801







[illegible]

Db	6595	AAATGCCGAACACACTCACCACACCGCTGTGTGGAGAGACCGCCCTCATCTGCAG	7013
RESULT 4			
LOCUS	BC033087		
DEFINITION	BC033087 Homo sapiens, similar to nuclear receptor co-repressor 2, clone IMAGE:5016291, mRNA.	2000 bp	linear HTC 04-MAR-2003
ACCESSION	BC033087		
VERSION	BC033087.1	GI:21619972	
KEYWORDS	HTC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 2000)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
COMMENT	Contact: MGC help desk Email: <a href="mailto:gcgabs-remail.nih.gov">gcgabs-remail.nih.gov</a> Tissue Procurement: DCTD/DRP cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada <a href="mailto:info@cgsc.bc.ca">info@cgsc.bc.ca</a> Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fiehl, Erin Garland, Ran Guin, Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeary, Steven Ness, Pavan Pandon, Anna-Liisa Prabhu, Parvaneh Saeei, Jacqueline Schein, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Stolt, Michael Thorne, Miranda Tsai, Natafja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.		
FEATURES	<p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Series: IRAL Plate: 43 Row: 1 Column: 21 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5454073 This clone has the following problem: retained intron.</p> <p>Location/Qualifiers</p> <p>1..2000</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="IMAGE:5016291"</p> <p>/tissue_type="lung, small cell carcinoma"</p> <p>/clone_id="NIH_MGC_7"</p> <p>/lab_host="DH10B-R"</p> <p>/note="Vector: pOT87"</p>		
ORIGIN	<p>Query Match 17.2%; Score 1473.6; DB 4; Length 2000; Best local Similarity 96.8%; Pred. No. 1.8e-265; Matches 1503; Conservative 0; Mismatches 49; Indels 0; Gaps 0;</p>		
Db	1	CATCTCGGGCTCCACACAGCTTGTGGACAGCGTGAAGGCGCACTGAGCCCGCTACCC	60
	272	CATTCTGGAGATCCACACAGGCTGTGGACAGCGTGAAGGCGCACTGAGCCCGCTACCC	331
Db	61	GCCTCAGAGCTTCTTACCAAGTGCAGATGCGCCGACGACACGAGCTGTGGGCTCT	120
	332	GCCTCAGAGCTTCTTACCAAGTGCAGATGCGCCGACGACACGAGCTGTGGGCTCT	391
Db	121	GAGATACAGACCACTTCCGCGACTATGCTCCACCTGTGTGCGCGGCTCATCATCA	180

Db 392 GAGATACAGACCACTCCGCGAATATGCTCCCACTGCGCCGCTTCATCATCA 451  
 Qy 181 GCCCAGCGCGAGGCGCTCCCTGCTGTGATTCAGCCCGGAATGAAGGCTCCA 240  
 Db 452 GCCCAGCGCGAGGCGCTCCCTGCTGTGATTCAGCCCGGAATGAAGGCTCCA 511  
 Qy 241 GAGCTTCACCTGCGCGCAGAGTCCCATATACCTGCGAGTGGGAAAGTCAGAGT 300  
 Db 512 GAGCTTCACCTGCGCGCAGAGTCCCATATACCTGCGAGTGGGAAAGTCAGAGT 571  
 Qy 301 GAGCTTCATTTGAAAGAGGCGCTCGGCTTATGAGTCTGCTGACCCCTGCTGCAAC 360  
 Db 572 GAGCTTCATTTGAAAGAGGCGCTCGGCTTATGAGTCTGCTGACCCCTGCTGCAAC 631  
 Qy 361 GTACCCCTGCTGAGCCAGCGGCGCAGCTGCGGAGTCTGAAAGACTTCAACAGAGCGTAG 420  
 Db 632 GTACCCCTGCTGAGCCAGCGGCGCAGCTGCGGAGTCTGAAAGACTTCAACAGAGCGTAG 691  
 Qy 421 CTTGACGGGCAAGCTGGAACCGGCTGTCTCCCGCAGCCCGCAGCACTGACCTGAGCT 480  
 Db 692 CTTGACGGGCAAGCTGGAACCGGCTGTCTCCCGCAGCCCGCAGCACTGACCTGAGCT 751  
 Qy 481 GAGCTGTGTCCGCGCAGCTGTCTCAAGAGAGCTGATCAAGAACCTGAGCCGCTGGA 540  
 Db 752 GAGCTGTGTCCGCGCAGCTGTCTCAAGAGAGCTGATCAAGAACCTGAGCCGCTGGA 811  
 Qy 541 CCGAGAGATCAACCATGATGAGAGAGAGATCTTAAGCTGTAAGAGAGAGAGAGAGAGT 600  
 Db 812 CCGAGAGATCAACCATGATGAGAGAGAGATCTTAAGCTGTAAGAGAGAGAGAGAGT 871  
 Qy 601 GAGAGAGAGAGCTGCCAGCCGCGCGAGCTGAGAGAGCCGCTGACCCGCGCCATCGA 660  
 Db 872 GAGAGAGAGAGCTGCCAGCCGCGCGAGCTGAGAGAGCCGCTGACCCGCGCCATCGA 931  
 Qy 661 GTTGAAGAGACCGCAGCTGTGTGATCATCTACAGAGAGAACCGGAAGAGCTGAGC 720  
 Db 932 GTTGAAGAGACCGCAGCTGTGTGATCATCTACAGAGAGAACCGGAAGAGCTGAGC 991  
 Qy 721 TGACATCGGATCTTGAAGAGCTGCGGCGCCAGAGTGGAGCTGCGCTGTAACAACGAGC 780  
 Db 992 TGACATCGGATCTTGAAGAGCTGCGGCGCCAGAGTGGAGCTGCGCTGTAACAACGAGC 1051  
 Qy 781 CTCGACACCCCGCAGTATCATGAGAACATCAAAATTAACCAAGCGATGCGAGAGAGCT 840  
 Db 1052 CTCGACACCCCGCAGTATCATGAGAACATCAAAATTAACCAAGCGATGCGAGAGAGCT 1111  
 Qy 841 AATCTTGTACTTGAAGAGAGAGATCAAGCTGCGGAAACAATGAGAGAGAGTTCGCCA 900  
 Db 1112 AATCTTGTACTTGAAGAGAGAGATCAAGCTGCGGAAACAATGAGAGAGAGTTCGCCA 1171  
 Qy 901 GCGCTATGACAGCTATGAGAGGCTTGAAGAGAGAGTGGAGCGATGGAAGAACACCC 960  
 Db 1172 GCGCTATGACAGCTATGAGAGGCTTGAAGAGAGAGTGGAGCGATGGAAGAACACCC 1231  
 Qy 961 GCGCGCGCGCGCCAGAGAGAGAGTGGCGAGTACTAGAGAGAGAGTTCCTGAGAT 1020  
 Db 1232 CCGGCGCGCGCGCCAGAGAGAGAGTGGCGAGTACTAGAGAGAGAGTTCCTGAGAT 1291  
 Qy 1021 CCGGAGAGAGCGCGAGCTGAGAGAGCGCATGACAGAGAGAGTGGCGAGCGGCGAGTGG 1080  
 Db 1292 CCGGAGAGAGCGCGAGCTGAGAGAGCGCATGACAGAGAGAGTGGCGAGCGGCGAGTGG 1351  
 Qy 1081 GCTGTTCATGTGCGCGCGCGCGAGAGAGAGCGAGTGTGAGAGATCATGATGAGTGGCTTC 1140  
 Db 1352 GCTGTTCATGTGCGCGCGCGCGAGAGAGAGCGAGTGTGAGAGATCATGATGAGTGGCTTC 1411  
 Qy 1141 AGAGCAGAGAGAGCTTGAAGAGAGAGTGGCGAGCTGCGCTGATCCCGCCATGCTGTA 1200  
 Db 1412 AGAGCAGAGAGAGCTTGAAGAGAGAGTGGCGAGCTGCGCTGATCCCGCCATGCTGTA 1471  
 Qy 1201 CGAGCTGAGCAGAGAGCGCATCAAGTTCATCAATGAAGAGGCTTATGCGCGAGCCCAT 1260  
 Db 1472 CGAGCTGAGCAGAGAGCGCATCAAGTTCATCAATGAAGAGGCTTATGCGCGAGCCCAT 1531

Qy 1261 GAAGTGTACAAAGACCGCAGGTATGAACATGTGAGTGAAGAGAGAGAGCTT 1320  
 Db 1532 GAAGTGTACAAAGACCGCAGGTATGAACATGTGAGTGAAGAGAGAGAGCTT 1591  
 Qy 1321 CCGGAGAGAGTTCATGAGAGATCCCAAGAACTTTGCGCTGATGATCATCTTCTGAGAG 1380  
 Db 1592 CCGGAGAGAGTTCATGAGAGATCCCAAGAACTTTGCGCTGATGATCATCTTCTGAGAG 1651  
 Qy 1381 GAAGACAGTGGCTGAGTGGCTCTTATTTACTACCTGAGTGAAGAGATGAAGCTTAA 1440  
 Db 1652 GAAGACAGTGGCTGAGTGGCTCTTATTTACTACCTGAGTGAAGAGATGAAGCTTAA 1711  
 Qy 1441 GAGCTGTGAGAGCGAGCTATCGCGCGCGCAAGAGCGCAGAGCAACAACAGCAGCA 1500  
 Db 1712 GAGCTGTGAGAGCGAGCTATCGCGCGCGCAAGAGCGCTGCTGCTTAAAGTTCCA 1771  
 Qy 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1552  
 Db 1772 GCAAAAAGATCAACCGCGGAGTAAAGGCTCCAGCCAGAGAGTTCCTG 1823

RESULT 5  
 BC020427 1362 bp mRNA linear HTC 03-JAN-2002  
 LOCUS  
 DEFINITION Homo sapiens, clone IMAGE:4179307, mRNA.  
 ACCESSION BC020427  
 VERSION BC020427.1 GI:18043012  
 KEYWORDS  
 SOURCE HTC.  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.

REFERENCE  
 AUTHORS Strausberg, R.  
 TITLE Direct Submision  
 JOURNAL Submitted (03-JAN-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,  
 USA

REMARK  
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
 Tissue Procurement: David N. Louis, M.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [villalombcm.tmc.edu](mailto:villalombcm.tmc.edu)  
 Villalon, D.K., Luna, R.A., Hale, S.M., Huiyk, S., Lu, X., Garcia,  
 A.M., Hollway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,  
 Muzny, D.M., Guarnante, P., Yoon, V., Kowls, C., Martin, R.,  
 Lawrence, S., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
 Series: IRAX Plate: 12 Row: 1 Column: 23  
 This clone has the following problem: no polyA-tail.

FEATURES  
 source 1..1362

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4179307"  
 /cissue\_type="Brain, anaplastic oligodendroglioma with  
 1p/19q loss"  
 /clone\_lib="NCI CGAP\_Brm67"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

ORIGIN



Db	19	CACCCACAGACGCTCAGGGGCAACCCCTGCCCCCCCCCTTACTCTTCTTCGAGGCCACAGC	78
Qy	6512	TGCCCCGCTCTGGAACCTCCGCGGCCCAACCCAGTGAACCTTCACTCCGCCCCCGGACCAT	65711
Db	79	TGCCCCGCTCTGGAACCTCCGCGGCCCAACCCAGTGAACCTTCACTCCGCCCCCGGACCAT	138
Qy	6572	GGTGTCCCCGGGCCCCGGTGGCTCCCCCCCCAAGCCGAAGGGGGCCAGAAGTCTTCAGAGCCAAAC	66313
Db	139	GGTGTCCCCGGGCCCCGGTGGCTCCCCCCCCAAGCCGAAGGGGGCCAGAAGTCTTCAGAGCCAAAC	198
Qy	6632	AAGAAGTCGGCTCTTGGGTGGTGGTGAAGACGGATATTGAACCTGTGTCCCAACCGGAGGGC	6691
Db	199	AAGAAGTCGGCTCTTGGGTGGTGGTGAAGACGGATATTGAACCTGTGTCCCAACCGGAGGGC	258
Qy	6692	ATGACGGAAGCCAGGGCACTCCCGAAGTCTGTGTACCCGCTGTGTACCGGAGTGGGAA	6751
Db	259	ATGACGGAAGCCAGGGCACTCCCGAAGTCTGTGTACCCGCTGTGTACCGGAGTGGGAA	318
Qy	6752	CAGACGGAGCCCAAGAGATGGGCTTCGAAGTCTCCAGGCAACACAGCCAGCCCGCACCC	68111
Db	319	CAGACGGAGCCCAAGAGATGGGCTTCGAAGTCTCCAGGCAACACAGCCAGCCCGCACCC	378
Qy	6812	TTCTTCAAGCAAGCTGAACCGAGAGCAACTCCGCAATGGTCAAGTCCAAAGCAAGAGATC	68711
Db	379	TTCTTCAAGCAAGCTGAACCGAGAGCAACTCCGCAATGGTCAAGTCCAAAGCAAGAGATC	438
Qy	6872	AACAAGAGCTGAACACCCACAACCCGGAATGAGCTGTGAATACAATATGAGCAGCCTGGG	69311
Db	439	AACAAGAGCTGAACACCCACAACCCGGAATGAGCTGTGAATACAATATGAGCAGCCTGGG	498
Qy	6932	ACGGAAGATCTTCAATATGCTCCGCCATCAACCGGAACAGGCTTTATGACTTATGAAGCCAG	6991
Db	499	ACGGAAGATCTTCAATATGCTCCGCCATCAACCGGAACAGGCTTTATGACTTATGAAGCCAG	558
Qy	6992	GGGGTGCAGGAACATGCCAGCAACCAATGGGGCTGGAGGGCCATTAATGAAGAAGCAATC	7051
Db	559	GGGGTGCAGGAACATGCCAGCAACCAATGGGGCTGGAGGGCCATTAATGAAGAAGCAATC	618
Qy	7052	ATGGGTAAATATGACCAAGTGGGAAAGAGTCCCGCGCTCAACCGCCAAATGCTTTTAAACCTT	7111
Db	619	ATGGGTAAATATGACCAAGTGGGAAAGAGTCCCGCGCTCAACCGCCAAATGCTTTTAAACCTT	678
Qy	7112	CTGAATGCAATGCTCCAGCTGCGCTGCTATGCCCCATTAACCGCTGCTGAACGAGAGT	7171
Db	679	CTGAATGCAATGCTCCAGCTGCGCTGCTATGCCCCATTAACCGCTGCTGAACGAGAGT	738
Qy	7172	GACCAACACATCAACTCGCCCAAGTGTGGCGGGGAGGACAAAGTCTCTGGAGAACCCAGC	7231
Db	739	GACCAACACATCAACTCGCCCAAGTGTGGCGGGGAGGACAAAGTCTCTGGAGAACCCAGC	798
Qy	7232	AGCCGAAAGGCCAAGTCCCGCGCCCCCGGGCTGTGGCATCTGGGAGAACCGGCCCAACCTCTGTC	7291
Db	799	AGCCGAAAGGCCAAGTCCCGCGCCCCCGGGCTGTGGCATCTGGGAGAACCGGCCCAACCTCTGTC	858
Qy	7292	TCCTCAAGTGAC-TGGAGGGAGACT-GCAACCGCCGGAACCGCTCAACAA---CCGCG	7346
Db	859	TCCTCAAGTGAC-TGGAGGGAGACTGGAAACCGCCGGAACCGCTCAACAAACCGGCTG	918
Qy	7347	TGTGGAGGAAGAGGCTCTGTTCGCAAGTT--CCAAGCATTTCCCTTACAAACCCCTGAT	7404
Db	919	GTTGGAGGAAGAGGCTCTGTTCGCAAGTTTCCCGGCATTTCCCTTACAAACCCCTGAT	978
Qy	7405	CATGCGGCTGCAAGCGGGGTGCAT-----GGCTTCCCAACCCCAACCGGGCTCCCGCG	7458
Db	979	CATGCGGCTTGAACGAGGGGGGTGTTCAATGGTTTTCCTCAACCCCAACCGGGCTTCCTCCG	1038
Qy	7459	GAGCAGCGCGGCCCTTCGCTGGACCCCAACAGCTGTGGACGAGGAGCCCAAG	7510
Db	1039	CGGGGAAAGCGGGGCCCTTCTCCGTGGCCCCCAACAGCTGTGGGAGCAAG	1090

QY 6903 AGCTGAATACATATATACAGCCAGCCTGGAGCGAGATCTTCAATATGCCCGCATCACCG 6962  
DB 361 AGCTGAATACATATATACAGCCAGCCTGGAGCGAGATCTTCAATATGCCCGCATCACCG 420  
QY 6963 GAACAGGCTTATGACCTTATAGAACCGCGCTGCAAGAAATATGCCAGACCAATATG 7022  
DB 421 GAACAGGCTTATGACCTTATAGAACCGCGCTGCAAGAAATATGCCAGACCAATATG 480  
QY 7023 GCGTGAAGGCTTATGACCTTATAGAACCGCGCTGCAAGAAATATGCCAGACCAATATG 7082  
DB 461 GCGTGAAGGCTTATGACCTTATAGAACCGCGCTGCAAGAAATATGCCAGACCAATATG 540  
QY 7083 CGCGCTGACAGCGCAATGCTTTTAAACCTCTGATATGCCAGTCCAGCCTGCGCTGCTA 7142  
DB 541 CGCGCTGACAGCGCAATGCTTTTAAACCTCTGATATGCCAGTCCAGCCTGCGCTGCTA 600  
QY 7143 TGCCCATTAACCGCTGCTGACAGGAGTGAACCACTACCTTCCAGTGGCGGCG 7202  
DB 601 TGCCCATTAACCGCTGCTGACAGGAGTGAACCACTACCTTCCAGTGGCGGCG 660  
QY 7203 GGAAGGCAAGGCTCTGAGCAGACCCAGAGCGGAAAGCAAGTCCCGCGCCGCGGCG 7262  
DB 661 GGAAGGCAAGGCTCTGAGCAGACCCAGAGCGGAAAGCAAGTCCCGCGCCGCGGCG 720  
QY 7263 TGGCATTTGGGAGACCGGCCACCTCTGTCTCTCAAGTGACTCGAGAGGAGACTGCAAC 7322  
DB 721 TGGCATTTGGGAGACCGGCCACCTCTGTCTCTCAAGTGACTCGAGAGGAGACTGCAAC 780  
QY 7323 GCGGAGCGCGCTGACCAACCGCGGTGTGGAGAGACAGGCGCTTCCGAGTTTCCAGCG 7382  
DB 781 GCGGAGCGCGCTGACCAACCGCGGTGTGGAGAGACAGGCGCTTCCGAGTTTCCAGCG 840  
QY 7383 CATTCCTCTACACCGCTGATCATGC-GGCTGACAGCGGAGT--GTATGAGCTTCCCGAC 7439  
DB 841 CATTCCTCTACACCGCTGATCATGCNNGTGACAGCGGAGTGTGATGTTTCCCGAC 900  
QY 7440 CCCCAACCGGCGCTCCCGCGGAGCG 7466  
DB 901 CCCCAACCGGCGCTCCCGCGGAG 927

RESULT 8  
B0557144 846 bp mRNA linear EST 16-SEP-2002  
LOCUS AGSCORRT.10253238 NIH MGC.109 Homo sapiens cDNA clone  
DEFINITION IMAGE:6584961 5', mRNA sequence.  
ACCESSION B0557144 GI:22907440  
VERSION B0557144.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 846)  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cga@bbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: Agencourt Bioscience Corporation (LLNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1CM2796 row: c column: 09  
High quality sequence stop: 719.  
Location/Qualifiers  
1..846  
Source /organism="Homo sapiens"

Query Match 9.8%; Score 841.8; DB 5; Length 846;  
Best Local Similarity 99.8%; Pred. No. 4.8e-147;  
Matches 843; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
ORIGIN  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6584961"  
/lab\_type="teratocarcinoma, cell line"  
/lib\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_109"  
/note="Organ: ovary; Vector: pOTB1; Site 1: EcoRI; Site 2:  
XhoI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."  
QY 6730 GCTGCTGACCGGAGTGGGAAACAGACGAGCCGACGAGATGGGCTTCAAGTCTCCAG 6789  
DB 1 GCTGCTGACCGGAGTGGGAAACAGACGAGCCGACGAGATGGGCTTCAAGTCTCCAG 60  
QY 6790 CAACACCGACCGCGCCAGGCTTCTTACGACGAGTGAACGAGACCACTCCGCGCATGT 6849  
DB 61 CAACACCGACCGCGCGCGGCTTCTTACGACGAGTGAACGAGACCACTCCGCGCATGT 120  
QY 6850 CAAGTCCAAAGAGCAAGATCAACAAGAGCTGAACCCACAAACCGAGATGAGCTGA 6909  
DB 121 CAAGTCCAAAGAGCAAGATCAACAAGAGCTGAACCCACAAACCGAGATGAGCTGA 180  
QY 6910 ATACATATACAGCAGCCTGGAGCGAGATCTTCAATATGCCGATCACCGAAACAG 6969  
DB 181 ATACATATACAGCAGCCTGGAGCGAGATCTTCAATATGCCGATCACCGAAACAG 240  
QY 6970 CCTTATGACTTATAGAACCGCGGTGACAGAACATGCCAGACCAATATGGGCTGGA 7029  
DB 241 CCTTATGACTTATAGAACCGCGGTGACAGAACATGCCAGACCAATATGGGCTGGA 300  
QY 7030 GGGCATTAATTGAAGGAGCACTCATAGGTAATATGACAGTGGAGAGATCCCGCGCT 7089  
DB 301 GGGCATTAATTGAAGGAGCACTCATAGGTAATATGACAGTGGAGAGATCCCGCGCT 360  
QY 7090 CAGCGCAATGCTTTTAAACCTCTGAAATGCCAGTCCAGCCTGCGGCTGATGCCAT 7149  
DB 361 CAGCGCAATGCTTTTAAACCTCTGAAATGCCAGTCCAGCCTGCGGCTGATGCCAT 420  
QY 7150 AACGCTGCTGACGAGCGAGATGACCACTACCTCTGCGCAGGTGGCGGGAAGGC 7209  
DB 421 AACGCTGCTGACGAGCGAGATGACCACTACCTCTGCGCAGGTGGCGGGAAGGC 480  
QY 7210 CAAGTCTCTGAGAGACCGAGAGCGGAAAGCCAAAGTCCCGGCGCGGCTTGGCATC 7269  
DB 481 CAAGTCTCTGAGAGACCGAGAGCGGAAAGCCAAAGTCCCGGCGCGGCTTGGCATC 540  
QY 7270 TGGGAGACCGGACCGCTTGTCTCTCTGATGACCTGCGAGGAGAGATGCAACCGCGGAG 7329  
DB 541 TGGGAGACCGGACCGCTTGTCTCTCTGATGACCTGCGAGGAGAGATGCAACCGCGGAG 600  
QY 7330 GCGGCTGACCAACCGCGGTGTGGAGAGACAGGCGCTGCTGCGCAGGTTTCCAGGCTTCCC 7389  
DB 601 GCGGCTGACCAACCGCGGTGTGGAGAGACAGGCGCTGCTGCGCAGGTTTCCAGGCTTCCC 660  
QY 7390 CTACAAACCGCTGATGAGCGGCTGACGAGGAGTGTATGAGTTTCCGACCCGACCGG 7449  
DB 661 CTACAAACCGCTGATGAGCGGCTGACGAGGAGTGTATGAGTTTCCGACCCGACCGG 720  
QY 7450 CTTCCCGCGGAGCAGCGGCGCTTGTGCTGAGCGCCGACCAACCGCTTGGAGAGAGAGCCAA 7509  
DB 721 CTTCCCGCGGAGCAGCGGCGCTTGTGCTGAGCGCCGACCAACCGCTTGGAGAGAGAGCCAA 780  
QY 7510 GCCACTGCTTGTCTGCGAGTACAGACACTCTCCGACAGCGAGTACTACAGAACGGGCG 7569

Db 781 GGCACGCTCTGTGTCAGTAGACAGACACTCTCCGACAGAGTAGACTCAGAAACAGCGCG 840  
Qy 7570 GGGGG 7574  
Db 841 GGGGG 845

RESULT 9  
LOCUS BM553310 1126 bp mRNA linear EST 20-FEB-2002  
DEFINITION AGENCOURT 6572642 NIH\_MGC\_41 Homo sapiens cDNA clone IMAGE:5467163  
5', mRNA sequence.  
ACCESSION BM553310  
VERSION BM553310.1 GI:18791930  
KEYWORDS EST.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1126)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-r@mail.nih.gov  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LNCM1968 row: 1 column: 12  
High quality sequence stop: 674.  
Location/Qualifiers  
1. 1126  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5467163"  
/tissue\_type="amelanotic melanoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_41"  
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene),  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

ORIGIN  
Query Match 9.5%; Score 815.6; DB 3; Length 1126;  
Best Local Similarity 99.5%; Pred. No. 4e-142;  
Matches 818; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 4369 GCTCAAGTACGACACGGGCGCTCCACACTGCTCCAAAAGACAGAGTAGCTCCCT 4428  
Db 1 GCTCAAGTACGACACGGGCGCTCCACACTGCTCCAAAAGACAGAGTAGCTCCCT 60  
Qy 4429 CATCGGACGCCCCGCGCGAGCTTCCACCCGTGCAACCCGCTGATGTATGCGCGACCC 4488  
Db 61 CATCGGACGCCCCGCGCGAGCTTCCACCCGTGCAACCCGCTGATGTATGCGCGACCC 120  
Qy 4489 CCGGGAGCTGGAACGCGCTGCTACAGAGAGAGCTGAAGAGCGGCGCAAGGCGCGACG 4548  
Db 121 CCGGGAGCTGGAACGCGCTGCTACAGAGAGAGCTGAAGAGCGGCGCAAGGCGCGACG 180  
Qy 4549 CAGCTCGGGGGGCTCCATTGCGCGCGCGCGCGCTCATTTGCTCGAGCTGGGTAAAGCC 4608

Db 181 CAGCTCGGGGGGCTCCATTGCGCGCGCGCGCGCGCTCATTTGTCCTGAGCTGGGCAAGCC 240  
Qy 4609 GCGGACAGAGCCCCCTGACCTATAGAGACCAAGGAGCACTTTGCGGCACTCCACG 4668  
Db 241 GCGGACAGAGCCCCCTGACCTATAGAGACCAAGGAGCACTTTGCGGCACTCCACG 300  
Qy 4669 AGTTTCGCCCCGTGACCATAGCGGAGGCCACGCGCGCTGACAGAGAGGAGCTTTGCT 4728  
Db 301 AGTTTCGCCCCGTGACCATAGCGGAGGCCACGCGCGCTGACAGAGAGGAGCTTTGCT 360  
Qy 4729 CAGCAAGGATCCGAGAGCCGAAAGCTGACGCTCTGATGATGATGATGATGATGATGAT 4788  
Db 361 CAGCAAGGATCCGAGAGCCGAAAGCTGACGCTCTGATGATGATGATGATGATGATGAT 420  
Qy 4789 GCAAGACACCGGCGCGAGACCAACCAACCCCATCTGCGCTTATGAGACACTGCTTG 4848  
Db 421 GCAAGACACCGGCGCGAGACCAACCAACCCCATCTGCGCTTATGAGACACTGCTTG 480  
Qy 4849 GCGCGTGAAGTGGCGTGAAGCTGATGACAGCAATCCCTTGAGCTTGACCCCACTTC 4908  
Db 481 GCGCGTGAAGTGGCGTGAAGCTGATGACAGCAATCCCTTGAGCTTGACCCCACTTC 540  
Qy 4909 CATACCCGCGGATCTCTTGAAGAGAGCGAGCGCGCTGATGACCTGCGCGACACTGGC 4968  
Db 541 CATACCCGCGGATCTCTTGAAGAGAGCGAGCGCGCTGATGACCTGCGCGACACTGGC 600  
Qy 4969 CCCCAACCCCACTTCCGCGACCTGATCCACCTTACCTTACCTGCGGCGTACCCGAGAC 5028  
Db 601 CCCCAACCCCACTTCCGCGACCTGATCCACCTTACCTTACCTGCGGCGTACCCGAGAC 660  
Qy 5029 GCGGCGCTGAGAAACCGGACAGACATCATGATGATGATGATGATGATGATGATGATGAT 5088  
Db 661 GCGGCGCTGAGAAACCGGACAGACATCATGATGATGATGATGATGATGATGATGATGAT 720  
Qy 5089 CCAACAACAGCGCCACCGGATGAGCGGAGCGAGCTGATGATGATGATGATGATGATGAT 5148  
Db 721 CCAACAACAGCGCGACCGGATGAGCGGAGCGAGCTGATGATGATGATGATGATGATGAT 780  
Qy 5149 CGAGTCTCTGCTGCGACCTCACTAGCGTGGGTCGCGGAGG 5190  
Db 781 CGAGTCTCTGCTGCGACCTCACTAGCGTGGGTCGCGGAGG 822

RESULT 10  
LOCUS B0690869 984 bp mRNA linear EST 15-JUL-2002  
DEFINITION AGENCOURT 8034173 NIH\_MGC\_110 Homo sapiens cDNA clone IMAGE:6208241  
5', mRNA sequence.  
ACCESSION B0690869  
VERSION B0690869.1 GI:21816185  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 984)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LNCM2365 row: 5 column: 18  
High quality sequence stop: 671.  
Location/Qualifiers  
1. 984

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6208241"
/tissue_type="ductal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_11b="NIH_MGC_110"
/notes="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

```

## ORIGIN

Query Match 9.3%; Score 798.2; DB 5; Length 984;  
 Best Local Similarity 94.7%; Pred. No. 7.3e-139;  
 Matches 892; Conservative 0; Mismatches 39; Indels 11; Gaps 6;

```

4001 GGCATCCGCGGAGACAGACAGCCCGCCACCTCAAGAGACAGACCACTCCGCGG 4060
1 GGCATCCGCGGAGACAGACAGCCCGCCACCTCAAGAGACAGACCACTCCGCGG 60
4061 TCCATCACAAGAGGATCCCTCGTCTACGTGAGGACACAGAGAGACTTACCTGCTCG 4120
61 TCCATCACAAGAGGATCCCTCGTCTACGTGAGGACACAGAGAGACTTACCTGCTCG 120
4121 GAGGCAAGCTCTTAAAGGGAGGAGGACAGCTCCGCGCCCGCCCTCAAGAGAGCTG 4180
121 GAGGCAAGCTCTTAAAGGGAGGAGGACAGCTCCGCGCCCGCCCTCAAGAGAGCTG 180
4181 ACCGAGGCTTACAAGACGACAGGCGCTGGGCGCCCTGAAAGCTGAAGCGGCGCATGAGGCG 4240
181 ACCGAGGCTTACAAGACGACAGGCGCTGGGCGCCCTGAAAGCTGAAGCGGCGCATGAGGCG 240
4241 CTGTGTGCGCAGCGTGAAGAGGCGGCGCTTCATTCATGAGATCCCGCGAGAGAGCTG 4300
241 CTGTGTGCGCAGCGTGAAGAGGCGGCGCTTCATTCATGAGATCCCGCGAGAGAGCTG 300
4301 CGGCAACGCGCCGACAGCTGCGCTTGGCGCGCGCGCTTCAAGAGAGGCTCCATCAGCAG 4360
301 CGGCAACGCGCCGACAGCTGCGCTTGGCGCGCGCGCTTCAAGAGAGGCTCCATCAGCAG 360
4361 GGCACCCCGCTCAAGTACGACACCGGCGGCTCAACACCTGCTCCAAAAAGACGACGTA 4420
361 GGCACCCCGCTCAAGTACGACACCGGCGGCTCAACACCTGCTCCAAAAAGACGACGTA 420
4421 CGCTCCCTCATGAGGAGCCCGCGGCGGAGCTTCCACCCGCTGATGATGATG 4480
421 CGCTCCCTCATGAGGAGCCCGCGGCGGAGCTTCCACCCGCTGATGATGATG 480
4481 GCGGACCGCGCGGAGCTGGAAGCTGCTTCAAGAGAGAGCTGAAGAGCGGCGCAAGG 4540
481 GCGGACCGCGCGGAGCTGGAAGCTGCTTCAAGAGAGAGCTGAAGAGCGGCGCAAGG 540
4541 ACCGCGCAG-CAGCTGCGGCGGCTGCTGAGGCGGCGCGGCTGATGATGCTGAGCT 4599
541 ACCGCGCAGCTGCTGAGGCGGCTGCTGAGGCGGCGGCTGATGATGCTGAGCT 600
4600 GGGTAAGCGCGGAGAGCCCGCTTGAAGAGAGCAACGCGGAGCAACCTTTGCGGCGCA 4659
601 GGGTAAGCGCGGAGAGCCCGCTTGAAGAGAGCAACGCGGAGCAACCTTTGCGGCGCA 660
4660 CCTCCACGAGGTTGGCCGCTGACCAAGGAGAGCGCGCGCTGACAGAGAGGAG 4719
661 CCTCCACGAGGTTGGCCGCTGACCAAGGAGAGCGCGCGCTGACAGAGAGGAG 720
4720 CTTTGTGTCAGCAAGGATCCCA-GGACCGAAGCTGAGCTGAGCGCTGTGAGATCG 4778
721 CTTTGTGTCAGCAAGGATCCCAAGGAGAGCGAAGCTGAGCTGAGCGCTGTGAGATCG 780

```

```

4779 CCAAGTCCCGGACAGCAGCTGTCGCGGAGACCAACCCACACCC--ATCTGCGCTATGA 4836
781 CCAAGTCCCGGACAGCAGCTGTCGCGGAGACCAACCCACACCCATCTCGCCCTATGA 840
4837 GCACCTGCTTCGCGGCGGAGTGGCGGTGAGAGCTGATTCGAGGAGC--ATCCCGCT 4890
841 GCACCTGCTTCGCGGCGGAGTGGCGGTGAGAGCTGATTCGAGGAGCATTCCCTG 900
4891 GGCCTTCGAGCCCACTTCATTA-CCCGCGGAGATCCCTCTG 4931
901 GGCCTTCGAGCCCACTTCATTAACCCCGGAGATCCCTCTG 942

```

RESULT 11  
 BM423558  
 LOCUS  
 DEFINITION BM423558 1051 bp mRNA linear EST 29-JAN-2002  
 AGENCOURT\_6399411 NIH\_MGC\_41 Homo sapiens cDNA clone IMAGE:5516745  
 5', mRNA sequence.  
 ACCESSION BM423558  
 VERSION BM423558.1 GI:18391770  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1051)  
 AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: DCTD/DRP  
 cDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNM)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNM at:  
<http://image.llnl.gov>  
 Plate: LLCM2017 row: n column: 10  
 High quality sequence atop: 667.

## FEATURES

## source

```

1..1051
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5516745"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

```

## ORIGIN

Query Match 9.3%; Score 796.4; DB 3; Length 1051;  
 Best Local Similarity 98.0%; Pred. No. 1.6e-138;  
 Matches 817; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

```

4369 GCTCAAGTACGACACCGGCGGTCCACCACT-GGCTCCAAAAAGAGACGTAAGCTCCC 4427
1 GCTCAAGTACGACCTCGCGGCTCCACCACTGGGCTCCAAAAAGACGACGTAAGCTCCC 60
4428 TCATCGGAGCGCCCGCGGAGCTTCCACCGGTCACCCGCTGGATGTGATGGCGCAG 4487
61 TCATCGGAGCGCCCGCGGAGCTTCCACCGGTCACCCGCTGGATGTGATGGCGCAG 120
4488 CCCGGGCACTGGAAGCTGCTTACGAGGAGCGCTGGAAGAGCGCGGCGCAGGAGCGGCA 4547

```



|||||  
Db 121 CCCGGGCACTGGAACGCTGCTTACGAGAGAGCCGGAAGACCGGCGGAGGAGCCGCA 180  
4548 GCAGCTCGGGGGGCTCTATTGCGCGGGGCGCCCGGCTCATTTGCTGAGCTGGTAAAC 4607  
181 GCAGCTCGGGGGGCTCTATTGCGCGGGGCGCCCGGCTCATTTGCTGAGCTGGGCAAGC 240  
4608 CGGGGAGAGCCCCCTGACCTTATGAGGAGCAAGGGGACCCCTTTGCGGCGCACCTCCAC 4667  
241 CGGGGAGAGCCCCCTTAACTTATGAGGAGCAAGGGGACCCCTTTGCGGCGCACCTCCAC 300  
4668 GAGGTTGCGCCGCTGACCATGCGGAGGAGCCACCGCGCGCTGACGAGAGGAGGCTTTGCT 4727  
301 GAGGTTGCGCCGCTGACCATGCGGAGGAGCCACCGCGCGCTGACGAGAGGAGGCTTTGCT 360  
4728 CCAGCAAGGATCCCGAGAGCCGAAAGTGAAGCTGAGAGGCTGCTGAGATGCCAATGCC 4787  
361 CCAGCAAGGATCCCGAGAGCCGAAAGTGAAGCTGAGAGGCTGCTGAGATGCCAATGCC 420  
4788 CGCAGAGCAGCCGTCGCGGAGAGCACCCACACCCCAATCTTGCCCTTATGAGCAGCTGCTTC 4847  
421 CGCAGAGCAGCCGTCGCGGAGAGCACCCACACCCCAATCTTGCCCTTATGAGCAGCTGCTTC 480  
4848 GGGGCGTGAAGTGGCTGAGACCTGATGACAGCAGATCCCGCTGAGCTTGCAGCCCACT 4907  
481 GGGGCGTGAAGTGGCTGAGACCTGATGACAGCAGATCCCGCTGAGCTTGCAGCCCACT 540  
4908 CCATACCCCGCGGAGCATCTCTGAGAGGAGCGCGCTGATCTGAGCCGAGACCTGAG 4967  
541 CCATACCCCGCGGAGCATCTCTGAGAGGAGCGCGCTGATCTGAGCCGAGACCTGAG 600  
4968 CCGCCGAAACCCATCTACCCGAGCAGCTGATCCCACTGATCCGCGGCTGAGCCGAGCA 5027  
601 CCGCCGAAACCCATCTACCCGAGCAGCTGATCCCACTGATCCGCGGCTGAGCCGAGCA 660  
5028 CCGCGCGCTGAGAGACCGGAGAGCAGCATCATCATGATGATGAGAGGCGCTTCCGCCC 5087  
661 CCGCGCGCTGAGAGACCGGAGAGCAGCATCATCATGATGATGAGAGGCGCTTCCGCCC 720  
5088 ACCACAAACGCGCAGCCGATGAGCGGAGAGCTGATGATGAGAGGCGCTTCCGCCC 5147  
721 ACCACAAACGCGCAGCCGATGAGCGGAGAGCTGATGATGAGAGGCGCTTCCGCCC 780  
QY 5148 GCGAGCTCTGCGTGGAGATCACTAGCTGCGGAGTCCCGAGGAGCATCATGAGCC 5201  
Db 781 GCGAGCTCTGCGTGGAGATCACTAGCTGCGGAGTCCCGAGGAGCATCATGAGCC 834

RESULT 12  
BM802749 1159 bp mRNA 1linear EST 05-MAR-2002  
AGENCY: NIH\_MGC 5459836 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:5575037  
5', mRNA sequence.  
BM802749  
BM802749.1 GI:19119572  
EST  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 1159)  
NIH-MGC <http://mgi.nci.nih.gov/>,  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph. D.  
Email: [gsabers-remail.nih.gov](mailto:gsabers-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL at:

<http://image.linl.gov>  
Plate: L1AM12324 row: k column: 06  
High quality sequence stop: 652.  
Location/Qualifiers  
1. 1159  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5575037"  
/issue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1ib="NIH\_MGC\_92"  
/note="Organ: testis; Vector: pCMV-SPORT6; Site: 1: NotI;  
Site 2: SalI; Cloned unidirectionally; 0.150-0.170 kb primed.  
Average insert size 2.5 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 9.3%; Score 796.2; DB 3; Length 1159;  
Best Local Similarity 94.4%; Pred. No. 1.8e-138;  
Matches 869; Conservative 0; Mismatches 48; Indels 4; Gaps 4;  
QY 4621 CCTGACCTATGAGAGCAGCGGAGCACCCTTGGCGGCACTTCCACGAGTTGCGCCGT 4680  
Db 1 CTTGACCTATGAGAGCAGCGGAGCACCCTTGGCGGCACTTCCACGAGTTGCGCCGT 60  
QY 4681 GACCATGCGGAGAGCCACCGCGCGCTGAGAGGAGGAGACCTTTGCTCCAGCAAGGATATC 4740  
Db 61 GACCATGCGGAGAGCCACCGCGCGCTGAGAGGAGGAGACCTTTGCTCCAGCAAGGATATC 120  
QY 4741 CCAGAGCGAAAGCTAGCGTGCAGCGCTGATGATGCGCAAGTCCCGCAGCAGCAGCT 4800  
Db 121 CCAGAGCGAAAGCTAGCGTGCAGCGCTGATGATGCGCAAGTCCCGCAGCAGCAGCT 180  
QY 4801 GCCCGAGACCAACCAACCCCACTGCGCTATGAGCACTTGCTTGGGGCGTGAAGTG 4860  
Db 181 GCCCGAGACCAACCAACCCCACTGCGCTATGAGCACTTGCTTGGGGCGTGAAGTG 240  
QY 4861 CGTGAGACTGTATGAGAGCAGCATATCCCGCTTGAGACCCCACTCCATATCCCGCGG 4920  
Db 241 CGTGAGACTGTATGAGAGCAGCATATCCCGCTTGAGACCCCACTCCATATCCCGCGG 300  
QY 4921 CATCCCTTGAAGCGAGCGGCTGCTACTACCTGAGCCGAGCACTTGAGCCGCGCAAC 4980  
Db 301 CATCCCTTGAAGCGAGCGGCTGCTACTACCTGAGCCGAGCACTTGAGCCGCGCAAC 360  
QY 4981 CTACCGGACCTGTACCAACCTTACTCATTCGCGGCTACCCCGAGAGCGGCGGTGGA 5040  
Db 361 CTACCGGACCTGTACCAACCTTACTCATTCGCGGCTACCCCGAGAGCGGCGGTGGA 420  
QY 5041 GAAACGCGAGCATATCATGATGATCATCATCATCTGAGAGATGACCAACACGCG 5100  
Db 421 GAAACGCGAGCATATCATGATGATCATCATCATCTGAGAGATGACCAACACGCG 480  
QY 5101 CACCGCATAGGCGCAGGAGCTGATATGCTGAGAGGAGCTCTGCGCCCGGAGAGTCTCGCT 5160  
Db 481 CACCGCATAGGCGCAGGAGCTGATATGCTGAGAGGAGCTCTGCGCCCGGAGAGTCTCGCT 540  
QY 5161 GGCATCTCACTACGCTGCGGAGTCCCGAGGAGCATATGATGATGATGATGATGATGAT 5220  
Db 541 GGCATCTCACTACGCTGCGGAGTCCCGAGGAGCATATGATGATGATGATGATGATGAT 600  
QY 5221 GCTGTGCTGCTGCGCCCGAGCAGCAGGAGCCGAGCAGCAGCAGCAGCAGCAGCAGCAG 5280  
Db 601 GCTGTGCTGCTGCGCCCGAGCAGCAGGAGCCGAGCAGCAGCAGCAGCAGCAGCAGCAG 660  
QY 5281 CTTCCCGACCGGCGCCAG-CCCTTACGAGCGGAGCAGAGCTCCCACTTCCCGAG 5339  
Db 661 CTTCCCGACCGGCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 720  
QY 5340 GAGGTCACACACTTGAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5399

Db 721 GAGGTCCACACACTTGACAAACCAACCAACGCTCTCTCCGAGCGGAGCAACC 780  
Qy 5400 GGGATCGAGAG-CGGGACCGGGATC-GGAGCGGGGAAAGTCATCTTCACCTCCAC-A 5456  
Db 781 GGAATCGAGAGCGGGACCGGGATCGGGGACCGGAAAGTCCTTCTTCACCTCCACCA 840  
Qy 5457 CGACGCTGAGACGACCCATCTGTGAGACCTGTGTCAGAGACGACGCGGACGACGG 5516  
Db 841 CAACGCTGGGACCCCGACCTCTGTGGAAACCTGTGAACCGAAGAAATATACCGGACAGCA 900  
Qy 5517 GCAGCAGCGCGGGGGTGGGG 5537  
Db 901 AGGGAACACACAGAGGGGG 921

RESULT 13  
BM477568 1057 bp mRNA linear EST 05-FEB-2002  
LOCUS AGENCOURT\_6484946 NIH\_MGC\_85 Homo sapiens cDNA clone IMAGE:5554420  
DEFINITION 5', mRNA sequence.  
ACCESSION BM477568  
VERSION BM477568.1 GI:18526610  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1057)  
NIH-MGC http://mgi.mci.nih.gov/.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: Lou Staudt  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LML)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LML, at:  
http://image.llnl.gov  
Plate: LMLM2272 row: p column: 05  
High quality sequence stop: 674.  
Location/Qualifiers  
1..1057  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5554420"  
/tissue\_type="lymphoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 85"  
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally, oligo-dT primed.  
Average insert size 1.867 Kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 9.3%; Score 795.2; DB 3; Length 1057;  
Best Local Similarity 93.5%; Pred. No. 2.7e-138;  
Matches 906; Conservative 0; Mismatches 29; Indels 34; Gaps 6;

Qy 2786 GCCGAGGGCGGCGAAGAAACCGGCTGTGTCCCAAGGCCAGGCTCTCAACCCGCACT 2845  
Db 1 GCCGAGGGCGGCGAAGAAACCGGCTGTGTCCCAAGGCCAGGCTCTCAACCCGCACT 60  
Qy 2846 GCGGACCCCGGGGCAATGCTCTACCCGAGGACCACTGAGCACTGAAGCACTGAAGCAG 2905  
Db 61 GCGGACCCCGGGGCAATGCTCTACCCGAGGACCACTGAGCACTGAAGCACTGAAGCAG 120  
Qy 2906 CGAGGGCTGCTATCCCCCATCCAGGTCAACCAAGTCCATGAGCCCCCGGAGAGAC 2965

Db 121 CGAGCGCTGCCATCCCGCCCATCCAGGTCAACAAAGTCATGAGCCCCCGGAGAGAC 180  
Qy 2966 GCAGCTCCCAACCAAGCAGCTCCCGCAGCCCGCAGCCGCAAAACCTGACGCGAG 3025  
Db 181 GCAGTCTCCCAACCAAGCAGCTCCCGCAGCCCGCAGCCGCAAAACCTGAGCGGAG 240  
Qy 3026 AGCGAGCCCTCAGACGCTTGGGAGCAGCCCCCGGGGCAAGAGCAGAGCCCGGACCC 3085  
Db 241 AGCGAGCCCTCAGACGCTTGGGAGCAGCCCCCGGGGCAAGAGCAGAGCCCGGACCC 300  
Qy 3086 CCGCGGCAAGAGG-----GCTTCGAGCGGAGGCCAG 3121  
Db 301 CCGCGGCAAGAGGAGCAGAGAGCTGTGTTCTTCCAGCTTTCGACGCGAGGCCAG 360  
Qy 3122 AAGCTGCTGGGGAACCCCTTGTGTAATTCCGCGCTGCTCCGTCGCCGCCCCCGT 3181  
Db 361 AAGCTGCTGGGGAACCCCTTGTGTAATTCCGCGCTGCTCCGTCGCCGCCCCCGT 420  
Qy 3182 GAGGTATCAAGGCTCCCGGATGCCCCGAGACCCCTCAGGCTTCTCTAGCTTCACTT 3241  
Db 421 GAGGTATCAAGGCTCCCGGATGCCCCGAGACCCCTCAGGCTTCTCTAGCTTCACTT 480  
Qy 3242 GGTACCCGACCTCGCGGCTCCATGACATGCGCGCGCGCTGCGCGGCGGCAACC 3301  
Db 481 GGTACCCGACCTCGCGGCTCCATGACATGCGCGCGCGCTGCGCGGCGGCAACC 540  
Qy 3302 AACATCTCAACCGCGCTCCCTCATCTCTCTGCAAGACCCGAGGCTCTGAGAG 3361  
Db 541 AACATCTCAACCGCGCTCCCTCATCTCTCTGCAAGACCCGAGGCTCTGAGAG 600  
Qy 3362 CAATAGGTGCTATCTCCCAAGAAATGCTGCTCAAGTCAAGTCCGTTACTCAGAGCAT 3421  
Db 601 CAATAGGTGCTATCTCCCAAGAAATGCTGCTCAAGTCCGTTACTCAGAGCAT 660  
Qy 3422 GCCAAGGCGCGGTTGGGCTGTCAACATGGGGGCTCCCGCCATGAGACCCCAAAAG 3481  
Db 661 GCCAAGGCGCGGTTGGGCTGTCAACATGGGGGCTCCCGCCATGAGACCCCAAAAG 720  
Qy 3482 CTGGCACCCTTCAGCGAGTGAAGAGAGAGAGTGTCCCAAGGCGCGGCTGAGGCA 3541  
Db 721 CTGGCACCCTTCAGCGAGTGAAGAGAGAGAGTGTCCCAAGGCGCGGCTGAGGCA 780  
Qy 3542 CCGGAGAGCTGGGGGTGCCCAAGGCCAGGAGGCTGCTGAGAGGAGAGGCTTGG 3601  
Db 781 CCGGAGAGCTGGGGGTGCCCAAGGCCAGGAGGCTGCTGAGAGGAGAGGCTTGG 838  
Qy 3602 GGTCTAGTTCCGGGCGGAGATCAACCAAGGCAATTC---AGACACGGGTGCCCTCG 3658  
Db 839 GGTCTAGTTCCGGGCGGAGATCAACCAAGGCAATTC---AGACACGGGTGCCCTCG 898  
Qy 3659 GACAGCG-CCATCAATACCGC-GGCTTCATACCCAGG--CAGCGCAGCTGACGTCC 3713  
Db 899 GACAGCGCATCAATACCGGAGTTTCATCAACCAAGGCAAGGCTGACGTCC 958  
Qy 3714 TGTACAGG 3722  
Db 959 TGTACAGG 967

RESULT 14  
BM910785 1022 bp mRNA linear EST 12-MAR-2002  
LOCUS AGENCOURT\_6617197 NIH\_MGC\_98 Homo sapiens cDNA clone IMAGE:5454337  
DEFINITION 5', mRNA sequence.  
ACCESSION BM910785  
VERSION BM910785.1 GI:19361164  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1022)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: L1CM1949 row: f column: 02  
 High quality sequence start: 12  
 High quality sequence stop: 554.

## FEATURES

source

1..1022  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:545437"  
 /issue\_type="astrocytoma grade IV, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_11b="NIH\_MGC\_98"  
 /note="Organ: Brain; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCACGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 9.2%; Score 787.6; DB 3; Length 1022;  
 Best Local Similarity 91.2%; Pred. No. 7.1e-137;  
 Matches 870; Conservative 0; Mismatches 79; Indels 5; Gaps 3;

6452 CACCCACAGCACTCAGCGCAGCCCGCCCGCCCTACTCTCCCTGGGGCAGC 6511  
 |||  
 13 CACCCACAGCACTCAGCGCAGCCCGCCCGCCCTACTCTCCCTGGGGCAGC 72  
 |||  
 6512 TCCCCGCTCTGAGACTCGCGCCGCCACCACTGACTCTACCTCCGCGCCGAGCAT 6571  
 |||  
 73 TCCCCGCTCTGAGACTCGCGCCGCCACCACTGACTCTACCTCCGCGCCGAGCAT 132  
 |||  
 6572 GGTGCCCCGCGCTGTGCTCTCCCAACAGGAGGGGCAAGAGTCTCCAGAGCCAAAC 6631  
 |||  
 133 GGTGCCCCGCGCTGTGCTCTCCCAACAGGAGGGGCAAGAGTCTCCAGAGCCAAAC 192  
 |||  
 6632 AAGAGGTGGTCTTGGTGTGTGATGAGCGGATTTGAACCTGTGTCGCCACCGAGGGC 6691  
 |||  
 193 AAGAGGTGGTCTTGGTGTGTGATGAGCGGATTTGAACCTGTGTCGCCACCGAGGGC 252  
 |||  
 6692 ATGACGAGCCAGGCACTCCGAGAGTGTGTATCCGCTCTGTAACCGGATGGGAA 6751  
 |||  
 253 ATGACGAGCCAGGCACTCCGAGAGTGTGTATCCGCTCTGTAACCGGATGGGAA 312  
 |||  
 6752 CAGACGAGCCAGGAGATGGGCTTCAGAGTTCAGAGGACAACACCGAGCGCCAGCC 6811  
 |||  
 313 CAGACGAGCCAGGAGATGGGCTTCAGAGTTCAGAGGACAACACCGAGCGCCAGCC 372  
 |||  
 6812 TTCTTCACGAGCTACCGAGAGCAATCCGCGATGTGTAATCCAGAAAGCAAGATC 6871  
 |||  
 373 TTCTTCACGAGCTACCGAGAGCAATCCGCGATGTGTAATCCAGAAAGCAAGATC 432  
 |||  
 6872 AACAGAACTTAACACCAACCGAGATGAGCTGAATATACATATCAGCAGCTGGG 6931  
 |||  
 433 AACAGAACTTAACACCAACCGAGATGAGCTGAATATACATATCAGCAGCTGGG 492  
 |||  
 6932 ACGGAGATTTCAATATGCGCGCATCACCGAAGACGGCTTATGACCTATAGAACCCAG 6991  
 |||  
 493 ACGGAGATTTCAATATGCGCGCATCACCGAAGACGGCTTATGACCTATAGAACCCAG 552

QY 6992 GCGGTGACAGAAATGCGACAGCAACATGGGGTGAAGCCATTAATGAAAGCACTC 7051  
 |||  
 DB 553 GCGGTGACAGAAATGCGACAGCAACATGGGGTGAAGCCATTAATGAAAGCACTC 612  
 |||  
 QY 7052 ATGGTAAATATGACAGTGGGAAGTCCCGCGCTCAGCGCCAAATGCTTTAAACCT 7111  
 |||  
 DB 613 ATGGTAAATATGACAGTGGGAAGTCCCGCGCTCAGCGCCAAATGCTTTAAACCT 672  
 |||  
 QY 7112 CTGAATGCGAGTGCAGAGCTGCGCGCTGTATAGCCATTAACCGCTCTACGAGACT 7171  
 |||  
 DB 673 CTGAATGCGAGTGCAGAGCTGCGCGCTGTATAGCCATTAACCGCTCTACGAGACT 732  
 |||  
 QY 7172 GACACACACTCACTCGCAGAGTGGCGGGAAGGACCAAGTCTTGGACAGACCCAGC 7231  
 |||  
 DB 733 GACACACACTCACTCGCAGAGTGGCGGGAAGGACCAAGTCTTGGACAGACCCAGC 792  
 |||  
 QY 7232 AGCCGAAAAGCCAG-TCGCCGCGCCCGGAGCTTGACATCTGGGA-CCGGACACCTTC 7287  
 |||  
 DB 793 AACCTAAAAGCCAGATCCCGCGCGCGCGGCTGATCTGGGGAACCGGCCACCTCC 852  
 |||  
 QY 7288 TGTCTCTCAATGACT-CGAGGAGACTGCAACCGCGGACGCGCTCACCAACCGCG 7346  
 |||  
 DB 853 TGTCTCTCAATGACTCGAGGAGAAATGACACCCCGGAAACCCGTTCAACCAAC 912  
 |||  
 QY 7347 TGTGAGAGGACAGGCGCTGTCGAGGTTCCAGCGCATTCCTCAACCCGCC 7400  
 |||  
 DB 913 CCGGAGTGGGAAAAAAGACCTTCGCGCGCGCGGCTCACACCCCGTTTCC 966  
 |||

## RESULT 15

BQ711119

875 bp mRNA linear EST 16-JUL-2002

## LOCUS

BQ711119 AGENCOURT\_7977236 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:6215773

## DEFINITION

5', mRNA sequence.

## ACCESSION

BQ711119 GI:21850018

## VERSION

EST.

## KEYWORDS

Homo sapiens (human)

## SOURCE

Homo sapiens

## ORGANISM

Homo sapiens

## REFERENCE

1 (bases 1 to 875)

## AUTHORS

NIH-MGC <http://mgc.nci.nih.gov/>

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: Dr. Mark Watson  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1CM2385 row: d column: 14  
 High quality sequence stop: 668.

## FEATURES

source

1..875  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6215773"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_11b="NIH\_MGC\_113"  
 /note="Organ: spleen; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCACGAG(G). Library constructed by Ling Hong in the  
 laboratory of Gerald M. Rubin (University of California,  
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
 Superscript II RT (Life Technologies). Note: this is a

## ORIGIN

NIH\_MGC Library."

Query Match 9.1%; Score 782; DB 5; Length 875;

Best Local Similarity 93.4%; Pred No. 7.9e-136; Matches 808; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

QY	6795	CCAGCCAGCCCGCAGCCTTCTTCAAGCAAGCTGACCGAGAGCAACTCCGCATGATCAAGT	6854
DB	1	CCAGCCAGCCCGCAGCCTTCTTCAAGCAAGCTGACCGAGAGCAACTCCGCATGATCAAGT	60
QY	6855	CCAAGAACGAGAGATCAACAAGAAAGCTGAACCCCAACCGGAATGAGCTGAATACA	6914
DB	61	CCAAGAACGAGAGATCAACAAGAAAGCTGAACCCCAACCGGAATGAGCTGAATACA	120
QY	6915	ATATCAGCAGCCTGGAGACGAGATCTTCAATATGCCCCCATCAACGGACCTTA	6974
DB	121	ATATCAGCAGCCTGGAGACGAGATCTTCAATATGCCCCCATCAACGGACCTTA	180
QY	6975	TGACCTATAGAAAGCAGGCGGTGACAGAAACATGCCAGCAACCAACATGGGGCTGAGGCCA	7034
DB	181	TGACCTATAGAAAGCAGGCGGTGACAGAAACATGCCAGCAACCAACATGGGGCTGAGGCCA	240
QY	7035	TAAATTAGAAAGGCACTCATGGTTAAATATGACCAATGGGAAAGATCCCCGCGCTCAGCG	7094
DB	241	TAAATTAGAAAGGCACTCATGGTTAAATATGACCAATGGGAAAGATCCCCGCGCTCAGCG	300
QY	7095	CCAATGCTTTTAAACCTCTGAATGCGAGTGCAGCCTGCCGCTGTATGCCCCATTAACCG	7154
DB	301	CCAATGCTTTTAAACCTCTGAATGCGAGTGCAGCCTGCCGCTGTATGCCCCATTAACCG	360
QY	7155	CTGCTGACGAGACGAGTGAACACACACTCACTGCGCAAGTGGCGGCGGAAAGCCAAAG	7214
DB	361	CTGCTGACGAGACGAGTGAACACACACTCACTGCGCAAGTGGCGGCGGAAAGCCAAAG	420
QY	7215	TCTCTGGAGAACCCGACGAGCCGAAAGCAAGTCCCCGGCCCCGGGCTTGGCATCTGGGG	7274
DB	421	TCTCTGGAGAACCCGACGAGCCGAAAGCAAGTCCCCGGCCCCGGGCTTGGCATCTGGGG	480
QY	7275	ACCGGCAACCCCTCTGTCTCTCTGCACTGCGAGGAGAGACTGCAACCGCCGGAAGCCGCG	7334
DB	481	ACCGGCAACCCCTCTGTCTCTCTGCACTGCGAGGAGAGACTGCAACCGCCGGAAGCCGCG	540
QY	7335	TCACCAACCGCGTGTGGAGAGACAGGCCCTCGTCCGCAAGTTCCAGCCATTCCCTTACA	7394
DB	541	TCACCAACCGCGTGTGGAGAGACAGGCCCTCGTCCGCAAGTTCCAGCCATTCCCTTACA	600
QY	7395	ACCCCTGATCATGCGGCTGACAGGCGGGTGTATGGCTTCCCAACCCCAACCGGGCTTCC	7454
DB	601	ACCCCTGATCATGCGGCTGACAGGCGGGTGTATGGCTTCCCAACCCCAACCGGGCTTCC	660
QY	7455	CCGCGGGAGAGCGGGCCCTCGCTGCGCCCCCAACAGCCTGGGACGAGAGCCCAAGCCAC	7514
DB	661	CCGCGGGAGAGCGGGCCCTCGCTGCGCCCCCAACAGCCTGGGACGAGAGCCCAAGCCAC	720
QY	7515	TGCTCTGCTCGAGTACGAGACACTTCTCGACAGCGAGTACTCAGAAACAGGGCGGGGGG	7574
DB	721	TGCTCTGCTCGAGTACGAGACACTTCTCGACAGCGAGTACTCA-AAAAGGCGNCGGGG	779
QY	7575	GGGCGGGCGGTEUBSPELUTCAAGTCCCAAGCAGCAAGAAACGGCCCTGCGAGAG	7634
DB	780	GGGCGGGGNNCCNN	839
QY	7635	CGGGGCGGCTGCCGACTCCCCCAAC	7659
DB	840	GGCGNCGCGCGCGCGCCCCCNC	864

Search completed: March 12, 2006, 19:33:44  
Job time : 19983 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Biocooperation Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2006, 04:54:20 ; Search time 3019 Seconds  
(without alignments)  
18921.193 Million cell updates/sec

Title: US-09-522-753-4  
Perfect score: 8564.8  
Sequence: 1 catgctgggctccacacagc.....caaaaaaaaaaaaaaaaaa 8571

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_21.\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*  
14: geneseqn2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8541	99.7	8561	12	ADG86298 Human SMR
2	8541	99.7	8561	12	ADN04304 Antipsoi
3	8495.2	99.2	8564	3	AAC74783 Human ORF
4	8459.2	98.8	8667	11	ACN44283 Human mRN
5	8423.4	98.3	8533	13	ADG84524 Human tum
6	8423.4	98.3	8533	13	ACN39603 Tumour-ab
7	8324.6	97.2	8686	8	ACA62249 Human enco
8	8324.6	97.2	8686	10	ADL13811 Osteoarth
9	8324.6	97.2	8686	12	ADG86290 Human SMR
10	8324.6	97.2	8686	12	ADG18820 Human bof
11	8226.8	96.1	9053	12	ADL12577 Human bfe
12	8222.2	96.0	9079	12	ADQ23294 Human bof
13	7554	75.4	7554	12	ADJ92815 Human co-
14	7337.6	85.7	7524	10	ADL13812 Osteoarth
15	7334.6	85.6	7521	8	ACA62250 Human nuc
16	5455	63.7	5989	6	ABK84305 Human CDN
17	5037.2	58.8	8544	8	ACA62451 Human enco
18	5037.2	58.8	8544	14	ADG61802 Murine NC
19	4797	56.0	7386	8	ACA62452 Mouse nuc

20	4543	53.0	7534	11	ACN44281	ACN44281 Mouse mRN
21	2618.6	30.6	2930	10	ADC35130	Adc35130 Human bre
22	2618.6	30.6	2930	12	ADG86301	Adg86301 Human SMR
23	1138.6	13.3	23380	11	ACN44282	ACN44282 Human gen
24	1132.4	13.2	220756	12	ADG86300	Adg86300 Human SMR
25	949.4	11.1	1027	3	AAC98781	Aac98781 Human pan
26	898.4	10.5	956	3	AAA98862	Aaa98862 Human pro
27	898.4	10.5	956	3	AAA98144	Aaa98144 Human pro
28	898.4	10.5	956	6	AAD29973	Aad29973 Human pha
29	760.6	8.9	2336	10	ADG81306	Adg81306 Human dia
30	652.8	7.6	752	10	ADL62646	Adl62646 Human apo
31	650	7.6	650	8	ACA57401	Aca57401 Human adl
32	630	7.4	7914	12	ADG83959	Adg83959 Human tum
33	630	7.4	7914	13	ADG87683	Adg87683 Human tum
34	630	7.4	7940	8	ABZ34833	Abz34833 Coding se
35	630	7.4	7940	13	ADG89791	Adg89791 Antagonis
36	630	7.4	7940	14	ADZ49335	Adz49335 Insulin B
37	629.6	7.4	7780	3	AAA60630	Aaa60630 HNRKR nuc
38	626.8	7.3	2745	12	ADJ92816	Adj92816 Human co-
39	626.8	7.3	2745	12	ADJ92816	Adj92816 Human HNR
40	555	6.5	555	8	ACA57524	Aca57524 Human adl
41	542.4	6.3	718	10	ADG76358	Adg76358 Human BSK
42	534	6.2	534	10	ADG79924	Adg79924 Leukaemia
43	494.6	5.8	2914	13	ADG07585	Adg07585 Full leng
44	458.8	5.4	527	13	ADU13876	Adu13876 Solid tum
45	432.8	5.1	91141	11	ACN44280	ACN44280 Mouse gen

## ALIGNMENTS

RESULT 1	ADG86298	standard; cDNA; 8561 BP.
ID	ADG86298	
AC	ADG86298	
XX		
DT	11-MAR-2004	(first entry)
XX		
DE	Human SMRT encoding cDNA SEQ ID NO:12.	
XX		
KW	SMRT, silencing mediator for retinoid and thyroid hormone action;	
KW	SMRT inhibitor; cytosolic; antiinflammatory; antiarthritic;	
KW	antirheumatic; antisense therapy; inflammatory disorder;	
KW	rheumatoid arthritis; hyperproliferative disorder; cancer; leukaemia;	
KW	breast cancer; human; gene; 88.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	2..7555
FT		/*tag= a
FT		/product= "SMRT"
XX		
PN	WO2003106645-A2.	
XX		
PD	24-DEC-2003.	
XX		
PF	17-JUN-2003; 2003WO-US018923.	
XX		
PR	17-JUN-2002; 2002US-00174014.	
XX		
PA	(ISIS-) ISIS PHARM INC.	
XX		
PI	Bennett CF, Freier SM, Doble KW;	
XX		
DR	WPI; 2004-082184/08.	
XX		
DR	P-PSDB; ADG86299.	
XX		
DR	GENBANK; NM_006312.	
XX		
PT	Novel antisense compound targeted to nucleic acid encoding SMRT	
PT	(silencing mediator for retinoid and thyroid hormone action), useful for	
PT	treating animal having disease associated with SMRT such as cancer,	

PT rheumatoid arthritis.  
XX  
PS Example 15; SEQ ID NO 12; 260bp; English.  
XX  
CC The present invention describes a compound (I) 8-50 nucleobases in length  
CC targeted to a nucleic acid molecule encoding SMRT (silencing mediator for  
CC retinoid and thyroid hormone action), where (I) specifically hybridises  
CC with the nucleic acid molecule encoding SMRT and inhibits expression of  
CC SMRT. (I) specifically hybridises with at least 8-nucleobase portion of a  
CC preferred target region on nucleic acid molecule encoding SMRT. Also  
CC described is a composition (II) comprising (I) and a carrier or diluent.  
CC (I) and (II) have cytostatic, antiinflammatory, antiarthritic and  
CC antitumoral activities, and can be used in antisense therapy, and as  
CC SMRT expression inhibitors. (II) is useful for inhibiting the expression  
CC of SMRT in cells or tissues. (I) is also useful for treating an animal  
CC having a disease or condition associated with SMRT, e.g., inflammatory  
CC disorder such as rheumatoid arthritis, or a hyperproliferative disorder  
CC such as cancer chosen from leukaemia and breast cancer, by inhibiting the  
CC expression of SMRT. (I) is useful for diagnostics, therapeutics,  
CC prophylaxis and as research reagents and kits. The present sequence  
CC encodes human SMRT, which is used in an example from the present  
CC invention. N.B. The present sequence is designated as SEQ ID NO:11 in  
CC example 15 but corresponds to SEQ ID NO:12 in the sequence listing.  
XX  
SQ Sequence 8561 BP; 1862 A; 3033 C; 2525 G; 1141 T; 0 U; 0 Other;  
Query Match 99.7%; Score 8541; DB 12; Length 8561;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 8560; Conservative 1; Mismatches 0; Indels 10; Gaps 1;  
QY 1 CATGTGCGGCTCCACACAGCTTTGTGGACAGACGTGAGAGGCGCACTAGAGCCCGCTACCC 60  
DB 1 CATGTGCGGCTCCACACAGCTTTGTGGACAGACGTGAGAGGCGCACTAGAGCCCGCTACCC 60  
QY 61 GCCCCACAGCCTTTTCTTACCCGATGCAATTCGCCCGGAGACACAGAGAGTGGGCTCT 120  
DB 61 GCCCCACAGCCTTTTCTTACCCGATGCAATTCGCCCGGAGACACAGAGAGTGGGCTCT 120  
QY 121 GGAGTACACGACCACTCCCGGACTATGCTCCCACTGTGCGCGGGCTCCATATCCA 180  
DB 121 GGAGTACACGACCACTCCCGGACTATGCTCCCACTGTGCGCGGGCTCCATATCCA 180  
QY 181 GCCCAGCGGCGAGAGGCTCTCCCTGCTGTCTGAGTTCCAGCCCGGAGATGAACGGTCCA 240  
DB 181 GCCCAGCGGCGAGAGGCTCTCCCTGCTGTCTGAGTTCCAGCCCGGAGATGAACGGTCCA 240  
QY 241 GGAGTCCACCTGCGGCGGACGATCCCACTCATATCTGCGCCGAGCTGGGAGATCAGAGAT 300  
DB 241 GGAGTCCACCTGCGGCGGACGATCCCACTCATATCTGCGCCGAGCTGGGAGATCAGAGAT 300  
QY 301 GGAGTTCAATTGAAACAAAGCGGCTCGGCTAGAGCTGCTGACCCCTGCTGCGACC 360  
DB 301 GGAGTTCAATTGAAACAAAGCGGCTCGGCTAGAGCTGCTGACCCCTGCTGCGACC 360  
QY 361 GTCACTCCCTGCTGCGCACGCGGCTGCGGATCTGAAAGACTCAACAAAGACCGTAG 420  
DB 361 GTCACTCCCTGCTGCGCACGCGGCTGCGGATCTGAAAGACTCAACAAAGACCGTAG 420  
QY 421 CCTGACGGGCAAGCTGGAACCGGTGTCTCCCGGAGCCCGGACACTGACCTCGAGCT 480  
DB 421 CCTGACGGGCAAGCTGGAACCGGTGTCTCCCGGAGCCCGGACACTGACCTCGAGCT 480  
QY 481 GGAGCTGTGCGCGCACGGCTGTCCAAAGGAGAGCTGATCCAGAAATATGACCGCGTGA 540  
DB 481 GGAGCTGTGCGCGCACGGCTGTCCAAAGGAGAGCTGATCCAGAAATATGACCGCGTGA 540  
QY 541 CCGAGAGATCACCATGTGTGAGACAGAGATCTTAAGCTGAAGAGACACAAACAGCT 600  
DB 541 CCGAGAGATCACCATGTGTGAGACAGAGATCTTAAGCTGAAGAGACACAAACAGCT 600  
QY 601 GGAGAGAGAGGCTGCGCAAGCCCGGAGGCTGGAAGCCCGGTGTCAACCGCCCATTCGA 660  
DB 601 GGAGAGAGAGGCTGCGCAAGCCCGGAGGCTGGAAGCCCGGTGTCAACCGCCCATTCGA 660

QY 661 GTCGAAGACCGCAGCCTGTGTCAGATCATACGACGAGAACCGGAAGAGCTGAAGC 720  
DB 661 GTCGAAGACCGCAGCCTGTGTCAGATCATACGACGAGAACCGGAAGAGCTGAAGC 720  
QY 721 TGCACATGGAATTTGGAAGGCTTGGGCGCCAGGTGAGCTGCGGTATACACAGCC 780  
DB 721 TGCACATGGAATTTGGAAGGCTTGGGCGCCAGGTGAGCTGCGGTATACACAGCC 780  
QY 781 CTCGACACCCGGGAGTATCATGAGACATCAAAATTAACGAGGATGCGGAAGGCT 840  
DB 781 CTCGACACCCGGGAGTATCATGAGACATCAAAATTAACGAGGATGCGGAAGGCT 840  
QY 841 AATCTTGATCTTCAAGAGAGGATCAAGCTCGAAGAACATGGAAGCAGAAATTTGCA 900  
DB 841 AATCTTGATCTTCAAGAGAGGATCAAGCTCGAAGAACATGGAAGCAGAAATTTGCA 900  
QY 901 GCGCTATGACCAAGCTCATAGAGGCTTGGAAAAGGAGCCCATCGAAGAACACC 960  
DB 901 GCGCTATGACCAAGCTCATAGAGGCTTGGAAAAGGAGCCCATCGAAGAACACC 960  
QY 961 GCGCGGCGGGCCAGAGAGAGCAAGTGCAGAGTACTACGAAGACATTCCTGAGAT 1020  
DB 961 GCGCGGCGGGCCAGAGAGAGCAAGTGCAGAGTACTACGAAGACATTCCTGAGAT 1020  
QY 1021 CCGCAAGACGCGCAGCTGCAAGAGCGCATGCAAGACAGGTTGGCCACGCGGCGAGTGG 1080  
DB 1021 CCGCAAGACGCGCAGCTGCAAGAGCGCATGCAAGACAGGTTGGCCACGCGGCGAGTGG 1080  
QY 1081 GCTGTCTATGTCGCGCGCCCGCAGCAGACAGAGGTTCAGAGATCATGATGGCTCTC 1140  
DB 1081 GCTGTCTATGTCGCGCGCCCGCAGCAGACAGAGGTTCAGAGATCATGATGGCTCTC 1140  
QY 1141 AGAGCAGAGAACTCGAGAAAGCAGATGCGCAGCTGACCGGATCCGCGCCATCTGTA 1200  
DB 1141 AGAGCAGAGAACTCGAGAAAGCAGATGCGCAGCTGACCGGATCCGCGCCATCTGTA 1200  
QY 1201 CGACGCTGACACGACCGCATCAAGTTTCATCAACATGAACGGGCTTATGGCCGACCCAT 1260  
DB 1201 CGACGCTGACACGACCGCATCAAGTTTCATCAACATGAACGGGCTTATGGCCGACCCAT 1260  
QY 1261 GAAAGGTATCAAAAGACCGCAGGTATGATGATGAGTGAACGAGAGAGAGACCTT 1320  
DB 1261 GAAAGGTATCAAAAGACCGCAGGTATGATGATGAGTGAACGAGAGAGAGACCTT 1320  
QY 1321 CCGGAGAAAGTTCAATGACGATCCCAAGAACTTGGCTGATCGCATATTCCTGGAAG 1380  
DB 1321 CCGGAGAAAGTTCAATGACGATCCCAAGAACTTGGCTGATCGCATATTCCTGGAAG 1380  
QY 1381 GAAAGCAAGTGGCTGAGTCCGCTCTTATTACTACCTGACTAAGAAATGAGAACTATTA 1440  
DB 1381 GAAAGCAAGTGGCTGAGTCCGCTCTTATTACTACCTGACTAAGAAATGAGAACTATTA 1440  
QY 1441 GAGCTGTGTGAGCGGAGCTATCGCGCGCGGCAAGGCCAGCAGCAACACAGACGA 1500  
DB 1441 GAGCTGTGTGAGCGGAGCTATCGCGCGCGGCAAGGCCAGCAGCAACACAGACGA 1500  
QY 1501 GCAGCAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAGAA 1560  
DB 1501 GCAGCAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGAGAA 1560  
QY 1561 AGATGAGAAAGAGAAAGAGGCGGAGAGAGAGAGAGAGAGAGGCGGAGGTGAGAA 1620  
DB 1561 AGATGAGAAAGAGAAAGAGGCGGAGAGAGAGAGAGAGAGAGGCGGAGGTGAGAA 1620  
QY 1621 CGACAAAGAAAGCTCTCTCAAGAGAGAAACAGACACCTTCAGGGGAGAGACAGAGGA 1680  
DB 1621 CGACAAAGAAAGCTCTCTCAAGAGAGAAACAGACACCTTCAGGGGAGAGACAGAGGA 1680  
QY 1681 GAAAGAGGCTGTGCGCTCCAAAGGCGCAAAATGCGCAACGCGGAGGAGACGCAAGG 1740  
DB 1681 GAAAGAGGCTGTGCGCTCCAAAGGCGCAAAATGCGCAACGCGGAGGAGACGCAAGG 1740

QY 1741 CCGCATCACCCGCTCATGTGCTAATGAGGCCAAACGGAGAGGCCATCAACCCCGCAGCA 1800  
DB 1741 CCGCATCACCCGCTCATGTGCTAATGAGGCCAAACGGAGAGGCCATCAACCCCGCAGCA 1800  
QY 1801 GAGCGCGAGCTGGCTCCATGGAAGTGAATGAGAGTTCTGCTGGAAGAGAGAAAT 1860  
DB 1801 GAGCGCGAGCTGGCTCCATGGAAGTGAATGAGAGTTCTGCTGGAAGAGAGAAAT 1860  
QY 1861 GGAACAGCCAAAGAAAGTCTCTGGAACAAGCGCGCAACTGGTCCGCGCATTCGCGCGAT 1920  
DB 1861 GGAACAGCCAAAGAAAGTCTCTGGAACAAGCGCGCAACTGGTCCGCGCATTCGCGCGAT 1920  
QY 1921 GGTGGGCTCCAGACGTGTGCGAGTGTAAAGACTTCTACTTCACTAACAGAGAGCA 1980  
DB 1921 GGTGGGCTCCAGACGTGTGCGAGTGTAAAGACTTCTACTTCACTAACAGAGAGCA 1980  
QY 1981 GAACTCTGATGAGATCTTGACAGACACAAGCTGAAGATGAGAGAGAGAGAAACGCGCG 2040  
DB 1981 GAACTCTGATGAGATCTTGACAGACACAAGCTGAAGATGAGAGAGAGAGAAACGCGCG 2040  
QY 2041 GAGGAAGAAAGAAAGCGCGCGCGCGCGCGCAGCGAGGAGGCTGCAATCCCGCGGTGT 2100  
DB 2041 GAGGAAGAAAGAAAGCGCGCGCGCGCGCGCAGCGAGGAGGCTGCAATCCCGCGGTGT 2100  
QY 2101 GAGGATGAGAGATGAGGCGCTCGGCGCTGAGCGGAATGAGAGAGAGATGCTGAGAGA 2160  
DB 2101 GAGGATGAGAGATGAGGCGCTCGGCGCTGAGCGGAATGAGAGAGAGATGCTGAGAGA 2160  
QY 2161 GGTGAAGCTTATCATGCTCTGGAATGAGGTGCCAGAGGGAATGCAGTGGCCAGC 2220  
DB 2161 GGTGAAGCTTATCATGCTCTGGAATGAGGTGCCAGAGGGAATGCAGTGGCCAGC 2220  
QY 2221 CACTGTCACACAGCTCAGACACCGAGAGATCCCTCTCTCTCACTGAGGCGCCCA 2280  
DB 2221 CACTGTCACACAGCTCAGACACCGAGAGATCCCTCTCTCTCACTGAGGCGCCCA 2280  
QY 2281 GGAACAAGGAGAGATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2340  
DB 2281 GGAACAAGGAGAGATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2340  
QY 2341 AAGGCCACCCACCCACACCGAGAGACATCCCGGCGCGCGCGCGCGCGCGCGCGCG 2400  
DB 2341 AAGGCCACCCACCCACACCGAGAGACATCCCGGCGCGCGCGCGCGCGCGCGCGCG 2400  
QY 2401 TGAAGGCAACCGAGCGCGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2460  
DB 2401 TGAAGGCAACCGAGCGCGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2460  
QY 2461 TGTGTCTCCCAAG 2520  
DB 2461 TGTGTCTCCCAAG 2520  
QY 2521 GAG 2580  
DB 2521 GAG 2580  
QY 2581 GCGCGCTCAAG 2640  
DB 2581 GCGCGCTCAAG 2640  
QY 2641 GAGAGCGCTGAG 2700  
DB 2641 GAGAGCGCTGAG 2700  
QY 2701 GCGCAGAGGCGACCATGCTGCAAGAGCTCGGCGCGCGCGCGCGCGCGCGCGCGCG 2760  
DB 2701 GCGCAGAGGCGACCATGCTGCAAGAGCTCGGCGCGCGCGCGCGCGCGCGCGCGCG 2760  
QY 2761 CTGCAAGTGAAG 2820  
DB 2761 CTGCAAGTGAAG 2820  
QY 2821 AAGGCCAAGCTCTTCAACCCCGACTGGGAGACCCCGGCGCAATGCTCAACCCAGAGCC 2880

DB 2821 AAGGCCAAGCTCTTCAACCCCGACTGGGAGACCCCGGCGCAATGCTCAACCCAGAGCC 2880  
QY 2881 ACTGACCTGGAAGACACTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
DB 2881 ACTGACCTGGAAGACACTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
QY 2941 AGTCAATGAGCGCGCGCGGAG 3000  
DB 2941 AGTCAATGAGCGCGCGCGGAG 3000  
QY 3001 GCGACGCAAAACCTGCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
DB 3001 GCGACGCAAAACCTGCAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
QY 3061 GGGCAAG 3120  
DB 3061 GGGCAAG 3120  
QY 3121 GAACTGTGCTGAG 3180  
DB 3121 GAACTGTGCTGAG 3180  
QY 3181 TGAAGTGAATCAAGGCGCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240  
DB 3181 TGAAGTGAATCAAGGCGCTCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240  
QY 3241 TGTGACCCACTGCGCGCTGAG 3300  
DB 3241 TGTGACCCACTGCGCGCTGAG 3300  
QY 3301 CACCATCTCAACCGCGCTCCCGCTCATCTCTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
DB 3301 CACCATCTCAACCGCGCTCCCGCTCATCTCTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
QY 3361 GCAAAATAGTGCATCTCCCAAG 3420  
DB 3361 GCAAAATAGTGCATCTCCCAAG 3420  
QY 3421 TGCAGAGGCGCGGCTGAG 3480  
DB 3421 TGCAGAGGCGCGGCTGAG 3480  
QY 3481 GCTGGACCGCTTCAAG 3540  
DB 3481 GCTGGACCGCTTCAAG 3540  
QY 3541 ACCGAGAGCGTGGGAG 3600  
DB 3541 ACCGAGAGCGTGGGAG 3600  
QY 3601 GGGCTGAGTTCCGGCGGAG 3660  
DB 3601 GGGCTGAGTTCCGGCGGAG 3660  
QY 3661 CAGCGCATCATATACCGGAG 3720  
DB 3661 CAGCGCATCATATACCGGAG 3720  
QY 3721 GGGCAACATCAACAGAGATCATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780  
DB 3721 GGGCAACATCAACAGAGATCATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780  
QY 3781 GGAACAGCTGCGCAAG 3840  
DB 3781 GGAACAGCTGCGCAAG 3840  
QY 3841 TGAAGGTGAGATGTCTGTAACCAAGTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3900  
DB 3841 TGAAGGTGAGATGTCTGTAACCAAGTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3900  
QY 3901 CCCCATGAG 3960



Db 3901 CCCCCATGAGACGCGCGCCCCCAAGCGCACTATGATCATGATGAGGCGCGGTGAG 3960  
Oy AGCCATCTCTCAGCGAGCATGAGAGGTCTCATYGGGCGGTGCGCATCCGCGGAGGACA 4020  
Db 3961 AGCCATCTCTCAGCGAGCATGAGAGGTCTCATYGGGCGGTGCGCATCCGCGGAGGACA 4020  
Oy 4021 CAGCCCCACACCTCAAGAGCAGACCAATCCGCGGCTCATCACAAGGATCCC 4080  
Db 4021 CAGCCCCACACCTCAAGAGCAGACCAATCCGCGGCTCATCACAAGGATCCC 4080  
Oy 4081 TCGGTCTTACGTGAGGAGCAGAGAGACTACTGTGTGAGGAGGCGCAAGCTCTTAAAGG 4140  
Db 4081 TCGGTCTTACGTGAGGAGCAGAGAGACTACTGTGTGAGGAGGCGCAAGCTCTTAAAGG 4140  
Oy 4141 GAGAGGACAGCTCTCGCGCCCCCAGCGCTCAAGGAGCTTGAAGGCTTAAAGGAGCA 4200  
Db 4141 GAGAGGACAGCTCTCGCGCCCCCAGCGCTCAAGGAGCTTGAAGGCTTAAAGGAGCA 4200  
Oy 4201 GGCCTTGGGCCCCCTGAAGCTGAAGCCGCGCCATGAGGGGCTGTGTGAGCAAGGAG 4260  
Db 4201 GGCCTTGGGCCCCCTGAAGCTGAAGCCGCGCCATGAGGGGCTGTGTGAGCAAGGAG 4260  
Oy 4261 GCGCGGCGGCTCTCATTCATGAGATCCCGGTGAGAGGCTGCGGCAACGCTCGAGCTGCC 4320  
Db 4261 GCGCGGCGGCTCTCATTCATGAGATCCCGGTGAGAGGCTGCGGCAACGCTCGAGCTGCC 4320  
Oy 4321 CCTGCGCCCGCGCGGCTCAAGAGAGGCTCATCAAGCAGGAGCCCGCTCAAGTACA 4380  
Db 4321 CCTGCGCCCGCGCGGCTCAAGAGAGGCTCATCAAGCAGGAGCCCGCTCAAGTACA 4380  
Oy 4381 CACCGGCGGCTCACACTGTGCTCAAAAAGCAGACGATCGCTCTTCATCGGAGCGCC 4440  
Db 4381 CACCGGCGGCTCACACTGTGCTCAAAAAGCAGACGATCGCTCTTCATCGGAGCGCC 4440  
Oy 4441 CGGCGGAGAGTTCACACCCGTGCAACCCGTGATGTGATGAGCGCGAGCGCGGAGCACTGA 4500  
Db 4441 CGGCGGAGAGTTCACACCCGTGCAACCCGTGATGTGATGAGCGCGAGCGCGGAGCACTGA 4500  
Oy 4501 AGGTGCTGTACGAGAGAGGCTGTAAGAGCGGCGAGGAGCCGCGAGAGCTTCGGGGGG 4560  
Db 4501 AGGTGCTGTACGAGAGAGGCTGTAAGAGCGGCGAGGAGCCGCGAGAGCTTCGGGGGG 4560  
Oy 4561 CTTCATTCGCGGCGCGCGCGCTCATGTGTGCTTGAAGTGGTAAAGCGCGGAGAGCGCC 4620  
Db 4561 CTTCATTCGCGGCGCGCGCGCTCATGTGTGCTTGAAGTGGTAAAGCGCGGAGAGCGCC 4620  
Oy 4621 CTTGACCTTATAGAGACACAGGGGACCTTTTGGCGGACCTTCCACAGAGTTCGCGCT 4680  
Db 4621 CTTGACCTTATAGAGACACAGGGGACCTTTTGGCGGACCTTCCACAGAGTTCGCGCT 4680  
Oy 4681 GACCATGCGGAGGCGCACGCGCGCTGAGAGGAGGAGCTTTTGTCCAGAGAGCATTC 4740  
Db 4681 GACCATGCGGAGGCGCACGCGCGCTGAGAGGAGGAGCTTTTGTCCAGAGAGCATTC 4740  
Oy 4741 CCAAGACCGAAAGCTGACGTGACGCTTGTGAGATTCGCAAGTCCCGGACAGACCGT 4800  
Db 4741 CCAAGACCGAAAGCTGACGTGACGCTTGTGAGATTCGCAAGTCCCGGACAGACCGT 4800  
Oy 4801 GCGCGAGACACACACACACACCTTGTGCTTATAGACACTGTTCGGGGGTGATGAGG 4860  
Db 4801 GCGCGAGACACACACACACACCTTGTGCTTATAGACACTGTTCGGGGGTGATGAGG 4860  
Oy 4861 CGTGAACCTGTATCGACAGCATATCCCTTGGCTTCAACCCCACTTCATATCCCGGAG 4920  
Db 4861 CGTGAACCTGTATCGACAGCATATCCCTTGGCTTCAACCCCACTTCATATCCCGGAG 4920  
Oy 4921 CATTCCTTGAACGAGCGGCTGCTTACTTACCTGCCCCGAGACACTGAGCCCAACCCAC 4980  
Db 4921 CATTCCTTGAACGAGCGGCTGCTTACTTACCTGCCCCGAGACACTGAGCCCAACCCAC 4980  
Oy 4981 CTATCCGACCTGTATCCACCTTACTTATCCGCGGCTTACCCGAGACAGGCGGCGCTGA 5040  
Db 4981 CTATCCGACCTGTATCCACCTTACTTATCCGCGGCTTACCCGAGACAGGCGGCGCTGA 5040

Oy 5041 GAACCGGAGACCATCATTAATGACTATACCTTCGAGGAGATGACCAACACGCGC 5100  
Db 5041 GAACCGGAGACCATCATTAATGACTATACCTTCGAGGAGATGACCAACACGCGC 5100  
Oy 5101 CACCGCATTCGCGGAGGAGCTGATATGCTGAGGGGCTTTCGCCCCGAGTCTCGCT 5160  
Db 5101 CACCGCATTCGCGGAGGAGCTGATATGCTGAGGGGCTTTCGCCCCGAGTCTCGCT 5160  
Oy 5161 GGCATCTCACTACGTGAGGGGCTCCCGAGGAGATCATGACCTGTCCAGAGTCCACACT 5220  
Db 5161 GGCATCTCACTACGTGAGGGGCTCCCGAGGAGATCATGACCTGTCCAGAGTCCACACT 5220  
Oy 5221 GCTGTGTCTGTGCCCCCGAGACACAGGAGACCCAGACACCGCATGAGACCGCTTGGCTA 5280  
Db 5221 GCTGTGTCTGTGCCCCCGAGACACAGGAGACCCAGACACCGCATGAGACCGCTTGGCTA 5280  
Oy 5281 CTTCCCAACCGCGCCCAAGCTTTCAGAGCGCGGACAGAGCTTCCCATCTTCCGAG 5340  
Db 5281 CTTCCCAACCGCGCCCAAGCTTTCAGAGCGCGGACAGAGCTTCCCATCTTCCGAG 5340  
Oy 5341 AGGTCCAAACACTTGAAGAAACCAACCAACGATCGTTCGAGCGGAGGAGAGACCG 5400  
Db 5341 AGGTCCAAACACTTGAAGAAACCAACCAACGATCGTTCGAGCGGAGGAGAGACCG 5400  
Oy 5401 GGATCGAGAGCGGAGCCGAGATCGGAGCGGAGAAAGTCCATCTCACTCCACACGAC 5460  
Db 5401 GGATCGAGAGCGGAGCCGAGATCGGAGCGGAGAAAGTCCATCTCACTCCACACGAC 5460  
Oy 5461 GGTGAGACACGACCCATCTGAGAGCTGTGACAGAGCAGACGCGGAGAGGAGCGAG 5520  
Db 5461 GGTGAGACACGACCCATCTGAGAGCTGTGACAGAGCAGACGCGGAGAGGAGCGAG 5520  
Oy 5521 CAGCGCGGGGGTGGGGGAGAGAGCGCGCGCGCTCCCATGACCATGACACAGCA 5580  
Db 5521 CAGCGCGGGGGTGGGGGAGAGAGCGCGCGCGCTCCCATGACCATGACACAGCA 5580  
Oy 5581 CTGCGCCATCTCCCTCGGAGCCAGAGATGCTTCAGAGAGACCCAGTGTGCTTACAA 5640  
Db 5581 CTGCGCCATCTCCCTCGGAGCCAGAGATGCTTCAGAGAGACCCAGTGTGCTTACAA 5640  
Oy 5641 CACAGGCAATGAAGGTATCATACCGCTGTGAGGCGCGAGCAAGGCCACAGTCTTAAAGTTC 5700  
Db 5641 CACAGGCAATGAAGGTATCATACCGCTGTGAGGCGCGAGCAAGGCCACAGTCTTAAAGTTC 5700  
Oy 5701 CACCTCCACCTTCACACCGTTCGCGGAGCTGCGCATTTCCACCTGACCCACCTGCC 5760  
Db 5701 CACCTCCACCTTCACACCGTTCGCGGAGCTGCGCATTTCCACCTGACCCACCTGCC 5760  
Oy 5761 ACTGGGCGGACCCCTGATGGGCTTACCTTACCTCATGAGAGCCGCTTGTGCTGCCAA 5820  
Db 5761 ACTGGGCGGACCCCTGATGGGCTTACCTTACCTCATGAGAGCCGCTTGTGCTGCCAA 5820  
Oy 5821 GAGAGCCCCCGGGGTGCGCCCGGACAGAGCGGCGCGAGACACCGGATGCTTCTCT 5880  
Db 5821 GAGAGCCCCCGGGGTGCGCCCGGACAGAGCGGCGCGAGACACCGGATGCTTCTCT 5880  
Oy 5881 GCGCAAGCCCCAGCGGCTCGGGCTGAGGCGGCGCTCTCCCGCAGAGAGGCTCGGA 5940  
Db 5881 GCGCAAGCCCCAGCGGCTCGGGCTGAGGCGGCGCTCTCCCGCAGAGAGGCTCGGA 5940  
Oy 5941 GCCCCGCGCCCTGATGCTCTGCTTCTGAGCAGCAGCATGCGCGCACCCCTGCGAA 6000  
Db 5941 GCCCCGCGCCCTGATGCTCTGCTTCTGAGCAGCAGCATGCGCGCACCCCTGCGAA 6000  
Oy 6001 GAACCTCGACCTTACACAGCGGAGCGGAGCCGCGGCGGACCTGCTCGGCTCGGA 6060  
Db 6001 GAACCTCGACCTTACACAGCGGAGCGGAGCCGCGGCGGACCTGCTCGGCTCGGA 6060  
Oy 6061 CCGGACCGGGAGAAAGACTCAAGTAAACCTTTTCACTCAGGAATGGAATCGGTTTC 6120  
Db 6061 CCGGACCGGGAGAAAGACTCAAGTAAACCTTTTCACTCAGGAATGGAATCGGTTTC 6120

QY 6121 TCTGGGTTACACGGGACAGCTACAGCCCGGAAAGGGGTGAAGCCCTCAGCCCTGTGAG 6180  
Db 6121 TCTGGGTTACACGGGACAGCTACAGCCCGGAAAGGGGTGAAGCCCTCAGCCCTGTGAG 6180  
QY 6181 CTCACCAAGTCTGACCCACGACAAAGGGGCTCCCAAGCACTTGAAAGAGTCCACAAAGAG 6240  
Db 6181 CTCACCAAGTCTGACCCACGACAAAGGGGCTCCCAAGCACTTGAAAGAGTCCACAAAGAG 6240  
QY 6241 CCAAGCTGAGGGGAGGTGGGCGCAAGGAGCCCGGTGAAGCTTGGCCGGGAGGC 6300  
Db 6241 CCAAGCTGAGGGGAGGTGGGCGCAAGGAGCCCGGTGAAGCTTGGCCGGGAGGC 6300  
QY 6301 CGGCCACTCCCAACCTGCGGCGCTGCTGAGAGCCAGCCCTGTCAGCCCGCTGCT 6360  
Db 6301 CGGCCACTCCCAACCTGCGGCGCTGCTGAGAGCCAGCCCTGTCAGCCCGCTGCT 6360  
QY 6361 CCAGACCGGCCCAAGGGGTCAAGAGGTCAAGCGGGGTGTCAACCTGAGCCAGCATCAG 6420  
Db 6361 CCAGACCGGCCCAAGGGGTCAAGAGGTCAAGCGGGGTGTCAACCTGAGCCAGCATCAG 6420  
QY 6421 TAGAGTCAACACAGGACTACACCGGGACCAACCAAGAGACTGAGCGCAACCCCTGCC 6480  
Db 6421 TAGAGTCAACACAGGACTACACCGGGACCAACCAAGAGACTGAGCGCAACCCCTGCC 6480  
QY 6481 CGCCCCCTCTACTCTCTTCCCTGAGGACAGCTGCCCTGCTGAGACTCTCGCGGCCAC 6540  
Db 6481 CGCCCCCTCTACTCTCTTCCCTGAGGACAGCTGCCCTGCTGAGACTCTCGCGGCCAC 6540  
QY 6541 CAGTGAACCTTCACTCTCCGCGCCGAGCAATGATGCCCCGCGGTGCTCCCCACAG 6600  
Db 6541 CAGTGAACCTTCACTCTCCGCGCCGAGCAATGATGCCCCGCGGTGCTCCCCACAG 6600  
QY 6601 CGAAGGGGGCAAGAGTCTTCAAGGCCAAAGAAGTCTGAGGTGTGTGAGGA 6660  
Db 6601 CGAAGGGGGCAAGAGTCTTCAAGGCCAAAGAAGTCTGAGGTGTGTGAGGA 6660  
QY 6661 CGGATTTGAACTGTGTCCCAACGGAGGATGACGAGCCAGAGGATGAGGAGTGC 6720  
Db 6661 CGGATTTGAACTGTGTCCCAACGGAGGATGACGAGCCAGAGGATGAGGAGTGC 6720  
QY 6721 TGTGTACCCGCTGTGTACCGGGATGAGGAAAGACGAGCCAGAGATGAGGAGTGC 6780  
Db 6721 TGTGTACCCGCTGTGTACCGGGATGAGGAAAGACGAGCCAGAGATGAGGAGTGC 6780  
QY 6781 GTTCTCAAGGCAACCAAGGAGCGGCGCAAGCTTCTTCAAGAGCTGACCGAGGCAATC 6840  
Db 6781 GTTCTCAAGGCAACCAAGGAGCGGCGCAAGCTTCTTCAAGAGCTGACCGAGGCAATC 6840  
QY 6841 CGCATGTCTCAAGTCAAGAGGAGAGATCAAGAAAGCTGAAACCCACAAACGGA 6900  
Db 6841 CGCATGTCTCAAGTCAAGAGGAGAGATCAAGAAAGCTGAAACCCACAAACGGA 6900  
QY 6901 TGAAGCTGAATACATATGAGCAGGCTGGGAGGAGATCTTCAATATGCCCCATC 6960  
Db 6901 TGAAGCTGAATACATATGAGCAGGCTGGGAGGAGATCTTCAATATGCCCCATC 6960  
QY 6961 CGAACAAGGCTTATGACCTTATGAAAGCCAGGCGGTGCAAGAAACATGCCACCAAT 7020  
Db 6961 CGAACAAGGCTTATGACCTTATGAAAGCCAGGCGGTGCAAGAAACATGCCACCAAT 7020  
QY 7021 GGGGCTGAGAGGCAATATGAAAGGCACTCATGGGTAAATATGACAGAGGGAAGTGC 7080  
Db 7021 GGGGCTGAGAGGCAATATGAAAGGCACTCATGGGTAAATATGACAGAGGGAAGTGC 7080  
QY 7081 CCGCGCGTCAAGGCGCAATGCTTTTAAACCTCTGAATGCGAGTGCAGGCTGCCTGC 7140  
Db 7081 CCGCGCGTCAAGGCGCAATGCTTTTAAACCTCTGAATGCGAGTGCAGGCTGCCTGC 7140  
QY 7141 TATGCCATTAACCGGTGCTGACGGAAGGAGTGAACACACTCACTTCGCGAGGTGC 7200  
Db 7141 TATGCCATTAACCGGTGCTGACGGAAGGAGTGAACACACTCACTTCGCGAGGTGC 7200  
QY 7201 CGGGAAGGCCAAGGTCTGTGGAGACCCAGCAGCCGAAAGCAAGTCCCGGCGCGG 7260

Db 7201 CGGGAAGGCCAAGGTCTGTGGAGACCCAGCAGCCGAAAGCAAGTCCCGGCGCGG 7260  
QY 7261 CTTGGCATCTGGGGAACCGGCAACCTCTGTCTCTCAAGTGAACCTCGAGGGAATCGCA 7320  
Db 7261 CTTGGCATCTGGGGAACCGGCAACCTCTGTCTCTCAAGTGAACCTCGAGGGAATCGCA 7320  
QY 7321 CGCGGGAAGCGGCTCAACCAACCGGTGTGAGAGGACAGGCGCTGTCGCGAGTTTCAC 7380  
Db 7321 CGCGGGAAGCGGCTCAACCAACCGGTGTGAGAGGACAGGCGCTGTCGCGAGTTTCAC 7380  
QY 7381 GCCATTCCTCTCAACCTCCCTGATCATGCGGCTGAGAGCGGGGTGTCAATGCTTCCAC 7440  
Db 7381 GCCATTCCTCTCAACCTCCCTGATCATGCGGCTGAGAGCGGGGTGTCAATGCTTCCAC 7440  
QY 7441 CCAACCGGAGCTTCCCGCGGAGAGCGGCGCTGCTGAGGCGCCCAACACGCTGGAGCA 7500  
Db 7441 CCAACCGGAGCTTCCCGCGGAGAGCGGCGCTGCTGAGGCGCCCAACACGCTGGAGCA 7500  
QY 7501 GAGGCCAAGCCACTGTGTGTGAGAGTACAGAGCACTTCCGACAGGAGTGTACAG 7560  
Db 7501 GAGGCCAAGCCACTGTGTGTGAGAGTACAGAGCACTTCCGACAGGAGTGTACAG 7560  
QY 7561 AACAGGCGGGGGGGCGGGCGGTGUBS.PEPLUGTCAAGTCCAGCGACCAAGAA 7620  
Db 7561 AACAGGCGGGGGGGCGGGCGGTGUBS.PEPLUGTCAAGTCCAGCGACCAAGAA 7620  
QY 7621 CGGCTCTGAGAGCGGGGGCGGCTGCCGACTCCCCCAACCAAGAAAGAGCCCTGAGT 7680  
Db 7621 CGGCTCTGAGAGCGGGGGCGGCTGCCGACTCCCCCAACCAAGAAAGAGCCCTGAGT 7680  
QY 7681 CGGCTGAGAGCGGGGGCGGCTGCCGACTCCCCCAACCAAGAAAGAGCCCTGAGT 7740  
Db 7681 CGGCTGAGAGCGGGGGCGGCTGCCGACTCCCCCAACCAAGAAAGAGCCCTGAGT 7740  
QY 7741 AACTAAGACTCCCGCCCGGCTGAGCTGTGAGACCTTACTCAGGGGATGTTTACTG 7800  
Db 7741 AACTAAGACTCCCGCCCGGCTGAGCTGTGAGACCTTACTCAGGGGATGTTTACTG 7800  
QY 7731 AACTAAGACTCCCGCCCGGCTGAGCTGTGAGACCTTACTCAGGGGATGTTTACTG 7790  
Db 7731 AACTAAGACTCCCGCCCGGCTGAGCTGTGAGACCTTACTCAGGGGATGTTTACTG 7790  
QY 7801 GTTCTCGGGAAGGAGGGAAGGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7860  
Db 7801 GTTCTCGGGAAGGAGGGAAGGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7860  
QY 7791 GTTCTCGGGAAGGAGGGAAGGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7850  
Db 7791 GTTCTCGGGAAGGAGGGAAGGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7850  
QY 7861 ACAAGGCGGCGAGGGCGGCGAGGAGCCCAAGCAGAGTACACGCACTTCCACGCA 7920  
Db 7861 ACAAGGCGGCGAGGGCGGCGAGGAGCCCAAGCAGAGTACACGCACTTCCACGCA 7920  
QY 7851 ACAAGGCGGCGAGGGCGGCGAGGAGCCCAAGCAGAGTACACGCACTTCCACGCA 7910  
Db 7851 ACAAGGCGGCGAGGGCGGCGAGGAGCCCAAGCAGAGTACACGCACTTCCACGCA 7910  
QY 7921 TGCCTCCCGGAATGCAATTTGGAACCAAGTCTTAACTGAGCTGCAAGCCCGCGCT 7980  
Db 7921 TGCCTCCCGGAATGCAATTTGGAACCAAGTCTTAACTGAGCTGCAAGCCCGCGCT 7980  
QY 7911 TGCCTCCCGGAATGCAATTTGGAACCAAGTCTTAACTGAGCTGCAAGCCCGCGCT 7970  
Db 7911 TGCCTCCCGGAATGCAATTTGGAACCAAGTCTTAACTGAGCTGCAAGCCCGCGCT 7970  
QY 7981 CCTTCGCGCTCCCATCCCGCTTACGCGCTGAGACATGAGCCGAGCCCTGTCAGGCC 8040  
Db 7981 CCTTCGCGCTCCCATCCCGCTTACGCGCTGAGACATGAGCCGAGCCCTGTCAGGCC 8040  
QY 7971 CCTTCGCGCTCCCATCCCGCTTACGCGCTGAGACATGAGCCGAGCCCTGTCAGGCC 8030  
Db 7971 CCTTCGCGCTCCCATCCCGCTTACGCGCTGAGACATGAGCCGAGCCCTGTCAGGCC 8030  
QY 8041 CCAAGTGGCTGCTGCTGCGGCTCCCAAGACTGCGCCGAGCAAGAGTGTGGAACCA 8100  
Db 8041 CCAAGTGGCTGCTGCTGCGGCTCCCAAGACTGCGCCGAGCAAGAGTGTGGAACCA 8100  
QY 8031 CCAAGTGGCTGCTGCTGCGGCTCCCAAGACTGCGCCGAGCAAGAGTGTGGAACCA 8090  
Db 8031 CCAAGTGGCTGCTGCTGCGGCTCCCAAGACTGCGCCGAGCAAGAGTGTGGAACCA 8090  
QY 8101 GTTCAAGGCGAGTGGGCGGCAAAAGGGCGAGTGCAGGCTGAGGGGAAACGAGTCTCGA 8160  
Db 8101 GTTCAAGGCGAGTGGGCGGCAAAAGGGCGAGTGCAGGCTGAGGGGAAACGAGTCTCGA 8160  
QY 8091 GTTCAAGGCGAGTGGGCGGCAAAAGGGCGAGTGCAGGCTGAGGGGAAACGAGTCTCGA 8150  
Db 8091 GTTCAAGGCGAGTGGGCGGCAAAAGGGCGAGTGCAGGCTGAGGGGAAACGAGTCTCGA 8150  
QY 8161 GGAAGTGAAGTGTGTTTATGACATGTTTGCAGAGCGGCTGAGGAAAGGAGATGTA 8220  
Db 8161 GGAAGTGAAGTGTGTTTATGACATGTTTGCAGAGCGGCTGAGGAAAGGAGATGTA 8220  
QY 8151 GGAAGTGAAGTGTGTTTATGACATGTTTGCAGAGCGGCTGAGGAAAGGAGATGTA 8210  
Db 8151 GGAAGTGAAGTGTGTTTATGACATGTTTGCAGAGCGGCTGAGGAAAGGAGATGTA 8210  
QY 8221 AATGATGTGTGTTTACAGGGGATATTTTGTATTAACCTTCAATGAATTAATCAGATGT 8280  
Db 8221 AATGATGTGTGTTTACAGGGGATATTTTGTATTAACCTTCAATGAATTAATCAGATGT 8280  
QY 8211 AATGATGTGTGTTTACAGGGGATATTTTGTATTAACCTTCAATGAATTAATCAGATGT 8270  
Db 8211 AATGATGTGTGTTTACAGGGGATATTTTGTATTAACCTTCAATGAATTAATCAGATGT 8270  
QY 8281 TTAAGCAAGAAAGACTTACCAAGTATTAATGCTGTGCTGCTTTGATCTTGCCTTACCG 8340  
Db 8281 TTAAGCAAGAAAGACTTACCAAGTATTAATGCTGTGCTGCTTTGATCTTGCCTTACCG 8340

[illegible]

D	b	6	1	GC	CC	CA	CA	G	C	T	T	T	C	T	A	C	C	A	G	T	G	C	A	G	A	T	G	C	C	C	G	A	C	A	C	G	A	C	G	T	C	C	T	C	T	120				
Q	y	121		GG	A	G	T	A	C	C	A	G	C	A	C	T	T	C	C	G	G	A	C	T	A	T	A	G	C	T	T	C	C	A	C	T	T	G	T	C	C	G	G	T	C	A	T	C	A	180
D	b	121		GG	A	G	T	A	C	C	A	G	C	A	C	T	T	C	C	G	G	A	C	T	A	T	A	G	C	T	T	C	C	A	C	T	T	G	T	C	C	G	G	T	C	A	T	C	A	180
Q	y	181		GC	CC	CA	CG	CG	GG	AG	GC	C	T	C	C	T	G	T	G	T	C	A	G	T	T	C	A	G	C	C	CG	GG	A	T	G	A	T	G	A	T	G	A	T	G	A	240				
D	b	181		GC	CC	CA	CG	CG	GG	AG	GC	C	T	C	C	T	G	T	G	T	C	A	G	T	T	C	A	G	C	C	CG	GG	A	T	G	A	T	G	A	T	G	A	T	G	A	240				
Q	y	241		GG	A	G	A	C	T	C	A	C	T	G	GG	CC	AG	T	C	C	A	T	C	A	C	T	T	G	C	C	CG	AG	T	C	C	A	G	T	G	GG	GG	A	T	G	A	300				
D	b	241		GG	A	G	A	C	T	C	A	C	T	G	GG	CC	AG	T	C	C	A	T	C	A	C	T	T	G	C	C	CG	AG	T	C	C	A	G	T	G	GG	GG	A	T	G	A	300				
Q	y	301		GG	A	G	T	T	C	A	T	T	G	A	A	G	C	A	G	C	C	T	C	G	T	A	G	A	G	C	T	G	A	G	C	T	G	A	G	C	C	T	T	G	A	360				
D	b	301		GG	A	G	T	T	C	A	T	T	G	A	A	G	C	A	G	C	C	T	C	G	T	A	G	A	G	C	T	G	A	G	C	T	G	A	G	C	C	T	T	G	A	360				
Q	y	361		GT	C	A	C	C	C	T	G	T	G	GC	A	C	GG	CC	AG	C	C	T	G	CG	A	T	T	G	A	A	C	T	G	A	A	C	T	G	A	A	C	C	T	G	A	420				
D	b	361		GT	C	A	C	C	C	T	G	T	G	GC	A	C	GG	CC	AG	C	C	T	G	CG	A	T	T	G	A	A	C	T	G	A	A	C	T	G	A	A	C	C	T	G	A	420				
Q	y	421		CT	T	A	C	CG	GG	C	A	G	T	G	GC	A	C	GG	CC	AG	C	C	T	G	CG	A	T	T	G	A	A	C	T	G	A	A	C	T	G	A	A	C	C	T	G	A	480			
D	b	421		CT	T	A	C	CG	GG	C	A	G	T	G	GC	A	C	GG	CC	AG	C	C	T	G	CG	A	T	T	G	A	A	C	T	G	A	A	C	T	G	A	A	C	C	T	G	A	480			
Q	y	481		GG	A	G	A	C	T	G	T	G	GC	A	C	GG	CC	AG	C	C	T	G	CG	T	A	G	A	G	A	G	C	T	G	A	A	C	T	G	A	A	C	C	T	G	A	540				
D	b	481		GG	A	G	A	C	T	G	T	G	GC	A	C	GG	CC	AG	C	C	T	G	CG	T	A	G	A	G	A	G	C	T	G	A	A	C	T	G	A	A	C	C	T	G	A	540				
Q	y	541		CC	G	A	G	A	T	C	A	T	G	T	A	G	A	G	C	A	G	A	T	T	C	T	A	G	T	G	A	A	G	A	A	G	A	G	C	A	A	C	A	G	A	600				
D	b	541		CC	G	A	G	A	T	C	A	T	G	T	A	G	A	G	C	A	G	A	T	T	C	T	A	G	T	G	A	A	G	A	A	G	A	A	G	C	A	A	C	A	G	A	600			
Q	y	601		GG	A	G	A	G	A	G	C	T	C	A	A	G	C	CG	CC	AG	A	G	C	T	G	A	A	G	C	C	CG	T	C	A	C	CG	CC	C	A	T	G	A	660							
D	b	601		GG	A	G	A	G	A	G	C	T	C	A	A	G	C	CG	CC	AG	A	G	C	T	G	A	A	G	C	C	CG	T	C	A	C	CG	CC	C	A	T	G	A	660							
Q	y	661		GT	C	A	G	A	C	C	CG	A	C	T	G	T	G	C	A	G	A	T	C	T	A	C	A	C	A	G	A	A	C	CG	A	A	G	A	G	C	T	G	A	720						
D	b	661		GT	C	A	G	A	C	C	CG	A	C	T	G	T	G	C	A	G	A	T	C	T	A	C	A	C	A	G	A	A	C	CG	A	A	G	A	G	C	T	G	A	720						
Q	y	721		TG	C	A	C	A	T	CG	A	T	TT	CG	A	A	GG	CC	CT	G	GG	GG	CC	CG	A	G	T	G	A	G	C	T	G	C	T	G	T	A	C	A	C	A	780							
D	b	721		TG	C	A	C	A	T	CG	A	T	TT	CG	A	A	GG	CC	CT	G	GG	GG	CC	CG	A	G	T	G	A	G	C	T	G	C	T	G	T	A	C	A	C	A	780							
Q	y	781		CT	CG	A	C	A	C	CG	A	G	T	A	T	C	A	T	G	A	A	C	A	T	A	A	A	A	C	A	A	G	C	G	A	G	C	G	A	A	G	A	840							
D	b	781		CT	CG	A	C	A	C	CG	A	G	T	A	T	C	A	T	G	A	A	C	A	T	A	A	A	A	C	A	A	G	C	G	A	G	C	G	A	A	G	A	840							
Q	y	841		AA	T	T	T	G	T	A	C	T	T	C	A	A	G	A	G	A	G	A	T	C	A	G	C	T	C	G	A	A	C	A	A	T	G	A	A	G	A	G	A	900						
D	b	841		AA	T	T	T	G	T	A	C	T	T	C	A	A	G	A	G	A	G	A	T	C	A	G	C	T	C	G	A	A	C	A	A	T	G	A	A	G	A	G	A	900						
Q	y	901		GC	GT	A	T	A	C	A	G	T	C	A	T	G	A	G	C	C	T	T	G	A	A	A	A	A	A	G	T	G	A	G	C	C	A	T	C	G	A	A	A	960						
D	b	901		GC	GT	A	T	A	C	A	G	T	C	A	T	G	A	G	C	C	T	T	G	A	A	A	A	A	A	A	G	T	G	A	G	C	C	A	T	C	G	A	A	960						
Q	y	961		GC	GT	A	T	A	C	A	G	T	C	A	T	G	A	G	C	C	T	T	G	A	A	A	A	A	A	A	G	T	G	A	G	C	C	A	T	C	G	A	1020							
D	b	961		GC	GT	A	T	A	C	A	G	T	C	A	T	G	A	G	C	C	T	T	G	A	A	A	A	A	A	A	G	T	G	A	G	C	C	A	T	C	G	A	1020							
Q	y	1021		CC	G	A	C	A	G	C	G	A	G	T	C	A	G	A	G	C	C	A	T	G	A	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	1080					
D	b	1021		CC	G	A	C	A	G	C	G	A	G	T	C	A	G	A	G	C	C	A	T	G	A	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	1080					
Q	y	1081		GC	TT	C	A	T	G	C	CG	CC	CG	C	A	G	C	C	A	G	A	C	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	1140						
D	b	1081		GC	TT	C	A	T	G	C	CG	CC	CG	C	A	G	C	C	A	G	A	C	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	G	A	1140						
Q	y	1141		AG	A	C	A	G	A	A	C	T	T	G	A	A	G	A	G	A	G	A	T	G	C	C	A	C	T	G	A	A	G	A	G	A	G	A	G	A	G	A	G	A	1200					
D	b	1141		AG	A	C	A	G	A	A	C	T	T	G	A	A	G	A	G	A	G	A	T	G	C	C	A	C	T	G	A	A	G	A	G	A	G	A	G	A	G	A	G	A	1200					

QY 1201 CGACGCTGACAGCAGCGCATCAAGTTCAATGAACGGGCTTATGGCCGACCCCAT 1260  
DB 1201 CGACGCTGACAGCAGCGCATCAAGTTCAATGAACGGGCTTATGGCCGACCCCAT 1260  
QY 1261 GAAAGGTGTACAAAGACCGCCAGGTCATGAAATGTGGAGTGAACAGAGAGAGACCTT 1320  
DB 1261 GAAAGGTGTACAAAGACCGCCAGGTCATGAAATGTGGAGTGAACAGAGAGAGACCTT 1320  
QY 1321 CCGGGGAAAGTTCAATCAGCATGCCAAGAACTTTGGCTTGATGGCAATTCCTTGAGAG 1380  
DB 1321 CCGGGGAAAGTTCAATCAGCATGCCAAGAACTTTGGCTTGATGGCAATTCCTTGAGAG 1380  
QY 1381 GAAGACATGGCTGAGTGGCTCTCTATTACTACTGACTAAGAAATGAGAACTATTA 1440  
DB 1381 GAAGACATGGCTGAGTGGCTCTCTATTACTACTGACTAAGAAATGAGAACTATTA 1440  
QY 1441 GAGCCTGTGTGAGACGGAGCTATCGGCGCGCGCAAGACCAAGCAACAGACGAC 1500  
DB 1441 GAGCCTGTGTGAGACGGAGCTATCGGCGCGCGCAAGACCAAGCAACAGACGAC 1500  
QY 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
DB 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
QY 1561 AGATGAG 1620  
DB 1561 AGATGAG 1620  
QY 1621 CGACAAAGAGAGACCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
DB 1621 CGACAAAGAGAGACCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
QY 1681 GAAAGAGGCTGTGGCTCTCCAAAGGCGCAAAATGCTCCCAAGCAGAGAGAGAGAG 1740  
DB 1681 GAAAGAGGCTGTGGCTCTCCAAAGGCGCAAAATGCTCCCAAGCAGAGAGAGAGAG 1740  
QY 1741 CGGCAATCAACCGCTCAATGGCTTAATGAGGCAACAGAGAGAGAGAGAGAGAGAG 1800  
DB 1741 CGGCAATCAACCGCTCAATGGCTTAATGAGGCAACAGAGAGAGAGAGAGAGAGAG 1800  
QY 1801 GAGCGCGAGCTGGCTCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860  
DB 1801 GAGCGCGAGCTGGCTCTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860  
QY 1861 GGAAGACGCGCAAGAAAGGTCTCTGAGAACAGCGGCGCAACTGGTCCGAGAT 1920  
DB 1861 GGAAGACGCGCAAGAAAGGTCTCTGAGAACAGCGGCGCAACTGGTCCGAGAT 1920  
QY 1921 GGTGGGCTCCAAAGCTGTGCGCAGGTAAAGAACTTCTAATTCAATCAAGAGAGAGAG 1980  
DB 1921 GGTGGGCTCCAAAGCTGTGCGCAGGTAAAGAACTTCTAATTCAATCAAGAGAGAGAG 1980  
QY 1981 GAACCTCGATGATGATTTTGCAGCAGCAGCAAGCTGAAGATGAGAGAGAGAGAGAG 2040  
DB 1981 GAACCTCGATGATGATTTTGCAGCAGCAGCAAGCTGAAGATGAGAGAGAGAGAGAG 2040  
QY 2041 GAGGAAAG 2100  
DB 2041 GAGGAAAG 2100  
QY 2101 GAGAGAGTGAAGATGAG 2160  
DB 2101 GAGAGAGTGAAGATGAG 2160  
QY 2161 GGCCTGAGGCTTATCATGCTCTGTGGAGATGAGTGGCCAGAGGGGAAATGCAAGTGGCC 2220  
DB 2161 GGCCTGAGGCTTATCATGCTCTGTGGAGATGAGTGGCCAGAGGGGAAATGCAAGTGGCC 2220  
QY 2221 CACTGTCAACACAGCTCAGACACGAGAGAGATCCCTCTCTCAACTGAGAGGCGCCAA 2280  
DB 2221 CACTGTCAACACAGCTCAGACACGAGAGAGATCCCTCTCTCAACTGAGAGGCGCCAA 2280

QY 2281 GGAACAAGGCGAGAAATGGGCGCAAGCCCGCAGCAGCCTTGTGGCGCGAGCGGCGCACCC 2340  
DB 2281 GGAACAAGGCGAGAAATGGGCGCAAGCCCGCAGCAGCCTTGTGGCGCGAGCGGCGCACCC 2340  
QY 2341 AGGCCCAACCCCAACCCCAACAGAGAGACATCCGGGGGCCCCATTGAGACCCACCCGGGCT 2400  
DB 2341 AGGCCCAACCCCAACCCCAACAGAGAGACATCCGGGGGCCCCATTGAGACCCACCCGGGCT 2400  
QY 2401 TGAAGCGACCGGAGGCCCTACGCCCCCAGCAGCAGCCCACTGCGCTCTGCACTCTCC 2460  
DB 2401 TGAAGCGACCGGAGGCCCTACGCCCCCAGCAGCAGCCCACTGCGCTCTGCACTCTCC 2460  
QY 2461 TGTGTCTCCCAAG 2520  
DB 2461 TGTGTCTCCCAAG 2520  
QY 2521 GAGAGAGCAG 2580  
DB 2521 GAGAGAGCAG 2580  
QY 2581 GCCCGTCAAG 2640  
DB 2581 GCCCGTCAAG 2640  
QY 2641 GAGAGCGGCTGAGGCGCACCGGAGGCGGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
DB 2641 GAGAGCGGCTGAGGCGCACCGGAGGCGGCTCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
QY 2701 CGGAGAGGCGCACCTGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760  
DB 2701 CGGAGAGGCGCACCTGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760  
QY 2761 CTGCAAGTGAAG 2820  
DB 2761 CTGCAAGTGAAG 2820  
QY 2821 AAGGCCCAAGCTCTTCAACCCCGACCTGGCGACCCCGGCGCAATGCTTCAACCCCAAGAG 2880  
DB 2821 AAGGCCCAAGCTCTTCAACCCCGACCTGGCGACCCCGGCGCAATGCTTCAACCCCAAGAG 2880  
QY 2881 ACTGAGCTGAAAG 2940  
DB 2881 ACTGAGCTGAAAG 2940  
QY 2941 AGTCCATGAGGCCCCCGGAG 3000  
DB 2941 AGTCCATGAGGCCCCCGGAG 3000  
QY 3001 GCCACCGCAAAACTTGCAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
DB 3001 GCCACCGCAAAACTTGCAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
QY 3061 GGGCAAG 3120  
DB 3061 GGGCAAG 3120  
QY 3121 GAAAGCTGCTGGAG 3180  
DB 3121 GAAAGCTGCTGGAG 3180  
QY 3181 TGAAGTATCAAGAGCTTCCCGCAGATGCCCGAGACCCCTCAGAGCTTCTCTTCAAGCTCAGC 3240  
DB 3181 TGAAGTATCAAGAGCTTCCCGCAGATGCCCGAGAGCCCTCAGAGCTTCTCTTCAAGCTCAGC 3240  
QY 3241 TGGTCAACCACTGCGCTTGGGCTCTCATGACATGCTGCGGCGGCTCTGCGCGCCACCC 3300  
DB 3241 TGGTCAACCACTGCGCTTGGGCTCTCATGACATGCTGCGGCGGCTCTGCGCGCCACCC 3300  
QY 3301 CACCATCTCAACCGGCGCTCCCTCATCTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
DB 3301 CACCATCTCAACCGGCGCTCCCTCATCTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
QY 3361 GCAAAATAGGTGCATCTCCCAAGAGATGTGGTCCAGCTTCAAGTCCGTAATCAGAGAGAG 3420

Db 3361 GCAAAATAGGTGCTCATCTCCCAAGGAATGCGGTCCAGCTCCAGCTCCCGTACTCCAGAGCA 3420  
Qy 3421 TGCCAAAGGCCCCGGGTGGGCCCCGTGTCAACATYGGGGCTGCCCTGTCCCATYGAACCCCAAAA 3480  
Db 3421 TGCCAAAGGCCCCGGGTGGGCCCCGTGTCAACATYGGGGCTGCCCTGTCCCATYGAACCCCAAAA 3480  
Qy 3481 GCTGAGCACTCTTACGCGAGTGAAGCAGAGCAGCTTCCCAAGGGCCAGGCTTGGGCC 3540  
Db 3481 GCTGAGCACTCTTACGCGAGTGAAGCAGAGCAGCTTCCCAAGGGCCAGGCTTGGGCC 3540  
Qy 3541 ACCGGAAGGCTTGGGGGTGCCCAAGGCCAGAGGCGTCCGTGTGAAGAGGACAGCTTCT 3600  
Db 3541 ACCGGAAGGCTTGGGGGTGCCCAAGGCCAGAGGCGTCCGTGTGAAGAGGACAGCTTCT 3600  
Qy 3601 GGGCTCACTTCCGGGCGGAAGATCAACAAAGGCAATTCACAGCAACGAGGTCCCTCGGA 3660  
Db 3601 GGGCTCACTTCCGGGCGGAAGATCAACAAAGGCAATTCACAGCAACGAGGTCCCTCGGA 3660  
Qy 3661 CAGCGCCATCACAATACCGGAGCTTCATCAACCAAGGCAAGCCAGCTGACGTCCTGTACA 3720  
Db 3661 CAGCGCCATCACAATACCGGAGCTTCATCAACCAAGGCAAGCCAGCTGACGTCCTGTACA 3720  
Qy 3721 GGGGCACTCAACAGATTCATGCGGAGGACAGCCCGAGTCCGCTTGGACCCGGGCGGGA 3780  
Db 3721 GGGGCACTCAACAGATTCATGCGGAGGACAGCCCGAGTCCGCTTGGACCCGGGCGGGA 3780  
Qy 3781 GGAACAGCTTCCCAAGGGCCAGTCTACAGAAAGGCAAGAGGCAAGTCTTGTCTTA 3840  
Db 3781 GGAACAGCTTCCCAAGGGCCAGTCTACAGAAAGGCAAGAGGCAAGTCTTGTCTTA 3840  
Qy 3841 TGAGGGTGGCATGTCTGTGAACCAAGTCTCAAGAGAGAGCGGCAAGAGCAGCTCAGAGCC 3900  
Db 3841 TGAGGGTGGCATGTCTGTGAACCAAGTCTCAAGAGAGAGCGGCAAGAGCAGCTCAGAGCC 3900  
Qy 3901 CCCCATGAGAGAGGCGGCCCCCAAGCGCACCTATGACATGATGAGAGGCGCGGTGGGCAAG 3960  
Db 3901 CCCCATGAGAGAGGCGGCCCCCAAGCGCACCTATGACATGATGAGAGGCGCGGTGGGCAAG 3960  
Qy 3961 AGCCATCTCTTACGAGCAGATCAGAGAGTCTCATYGGGCGTGCATCCGCGGAGAGCA 4020  
Db 3961 AGCCATCTCTTACGAGCAGATCAGAGAGTCTCATYGGGCGTGCATCCGCGGAGAGCA 4020  
Qy 4021 CAGCCCCACCACTTCAAAAGCAGACCAATCCGCGGCTCATCAACAAAGGATCCC 4080  
Db 4021 CAGCCCCACCACTTCAAAAGCAGACCAATCCGCGGCTCATCAACAAAGGATCCC 4080  
Qy 4081 TCGGTCCTACGTGAGAGGACAGAGAGACTTACCTTGGGTGGGAGGCCAAAGCTCTTAAAGCG 4140  
Db 4081 TCGGTCCTACGTGAGAGGACAGAGAGACTTACCTTGGGTGGGAGGCCAAAGCTCTTAAAGCG 4140  
Qy 4141 GAGAGGCAAGCTCCGCCCCCAAGCCCTCAAGGACCTGACCGAGGCTTACAAGACGA 4200  
Db 4141 GAGAGGCAAGCTCCGCCCCCAAGCCCTCAAGGACCTGACCGAGGCTTACAAGACGA 4200  
Qy 4201 GGCCTTGGGCCCCCTGAAGCTGAAGCGGAGCCATGAGGGCTGTGTGGCCACGCTGAAGGA 4260  
Db 4201 GGCCTTGGGCCCCCTGAAGCTGAAGCGGAGCCATGAGGGCTGTGTGGCCACGCTGAAGGA 4260  
Qy 4261 GGGGGGCGGCTTCATTCATGAGATCCCGGCGAGGAGCTGCGGACACGCCCGAGACTGCG 4320  
Db 4261 GGGGGGCGGCTTCATTCATGAGATCCCGGCGAGGAGCTGCGGACACGCCCGAGACTGCG 4320  
Qy 4321 CTTGAGCCCGCGGCTCAAGAGAGGCTCATCAAGCAGGAGCAACCCGCTCAAGTACGA 4380  
Db 4321 CTTGAGCCCGCGGCTCAAGAGAGGCTCATCAAGCAGGAGCAACCCGCTCAAGTACGA 4380  
Qy 4381 CACCGGCGGCTCACCACTGGCTCCAAAAGCAGACGATCGCTCATCGGAGGCC 4440  
Db 4381 CACCGGCGGCTCACCACTGGCTCCAAAAGCAGACGATCGCTCATCGGAGGCC 4440  
Qy 4441 CCGCGGAGAGTTCCACCCGTCACCCGCTGAGATGTGATGCGGACAGCCCGGAGCACTGGA 4500  
Db 4441 CCGCGGAGAGTTCCACCCGTCACCCGCTGAGATGTGATGCGGACAGCCCGGAGCACTGGA 4500

Db 4441 CCGCGGAGAGTTCCACCCGTCACCCGCTGAGATGTGATGCGGACAGCCCGGAGCACTGGA 4500  
Qy 4501 AGTGTCTGTACAGAGAGAGCTGAAAGCGCGGCAAGGACCCGCAAGCTTCCGGGGG 4560  
Db 4501 AGTGTCTGTACAGAGAGAGCTGAAAGCGCGGCAAGGACCCGCAAGCTTCCGGGGG 4560  
Qy 4561 CTTCATTCGCGGCGGCGCGGCTGATTTGTGCTGAGCTTGGGTAAAGCCCGGAGAGCC 4620  
Db 4561 CTTCATTCGCGGCGGCGCGGCTGATTTGTGCTGAGCTTGGGTAAAGCCCGGAGAGCC 4620  
Qy 4621 CTTGACCTATGAGAGCAAGGGGCAACCTTTTGGGCGGCAACCTTCCAGAGGTTCGCCGT 4680  
Db 4621 CTTGACCTATGAGAGCAAGGGGCAACCTTTTGGGCGGCAACCTTCCAGAGGTTCGCCGT 4680  
Qy 4681 GACCATGCGGAGAGCCCAAGCGGCGCTGACAGAGGAGGAGCCCTTTGTCCAGCAAGCATC 4740  
Db 4681 GACCATGCGGAGAGCCCAAGCGGCGCTGACAGAGGAGGAGCCCTTTGTCCAGCAAGCATC 4740  
Qy 4741 CCAAGGACGAAAGCTGACGCTGAGCGCTTGTGAGATTCGCAAGTCCCGCACAGACCTT 4800  
Db 4741 CCAAGGACGAAAGCTGACGCTGAGCGCTTGTGAGATTCGCAAGTCCCGCACAGACCTT 4800  
Qy 4801 GCGGAGGACCAACCCCAACCCCATCTGCGCCCTATGAGAGACCTTTCGGGGGTGAGTGG 4860  
Db 4801 GCGGAGGACCAACCCCAACCCCATCTGCGCCCTATGAGAGACCTTTCGGGGGTGAGTGG 4860  
Qy 4861 GCTGAGCCTGTATGCAAGCAACATCCCGCTTGGCTTCAAGCCCACTTCATACCCCGGG 4920  
Db 4861 GCTGAGCCTGTATGCAAGCAACATCCCGCTTGGCTTCAAGCCCACTTCATACCCCGGG 4920  
Qy 4921 CATCCCTTGGAGCGGAGCGGCTGCTTACTTACCTGCGCGGACACCTTGGCCCCCAACCCAC 4980  
Db 4921 CATCCCTTGGAGCGGAGCGGCTGCTTACTTACCTGCGCGGACACCTTGGCCCCCAACCCAC 4980  
Qy 4981 CTACCCGACCGCTGTACCCACCTTACCTCATCGCGGCTTACCCCGGACAGCGGCGCTGGA 5040  
Db 4981 CTACCCGACCGCTGTACCCACCTTACCTCATCGCGGCTTACCCCGGACAGCGGCGCTGGA 5040  
Qy 5041 GAAACCGGACGACCATCATCAATGAATGATCATGACCTTGGAGAGATGACCAACAGGAC 5100  
Db 5041 GAAACCGGACGACCATCATCAATGAATGATCATGACCTTGGAGAGATGACCAACAGGAC 5100  
Qy 5101 CACCGCATYGGCCCAAGCGAGCTGATATCTGAGGGGCTCTTCCGCGGAGTCTTCCCT 5160  
Db 5101 CACCGCATYGGCCCAAGCGAGCTGATATCTGAGGGGCTCTTCCGCGGAGTCTTCCCT 5160  
Qy 5161 GGCATCTCAACTACGTGAGGGGCCCCGAGGACATTCGACCTGTGACCTTCAAGTGCACACT 5220  
Db 5161 GGCATCTCAACTACGTGAGGGGCCCCGAGGACATTCGACCTGTGACCTTCAAGTGCACACT 5220  
Qy 5221 GCTGTGTCTGTGCCCCCGGACCAAGGACCCCAAGCCCAAGCATGGAACGCTTGTCTTA 5280  
Db 5221 GCTGTGTCTGTGCCCCCGGACCAAGGACCCCAAGCCCAAGCATGGAACGCTTGTCTTA 5280  
Qy 5281 CTTCCCAACCGGCCCCAGCCCTTACGACAGCGGCTCAAGAGCTTCCCATCTTCCCAAG 5340  
Db 5281 CTTCCCAACCGGCCCCAGCCCTTACGACAGCGGCTCAAGAGCTTCCCATCTTCCCAAG 5340  
Qy 5341 AGGTCCAAACAATTGAACAAACCAACCAACGCTCGGTCCAGGGGGAGGAGGAGCG 5400  
Db 5341 AGGTCCAAACAATTGAACAAACCAACCAACGCTCGGTCCAGGGGGAGGAGGAGCG 5400  
Qy 5401 GGATGAGAGGCGGAGCCGGGATCGGAGCGGAGAAAGTCCATTCCTCACTGACCCACGAC 5460  
Db 5401 GGATGAGAGGCGGAGCCGGGATCGGAGCGGAGAAAGTCCATTCCTCACTGACCCACGAC 5460  
Qy 5461 GGTGAGACACGACCCATCTGAGAGCTGTGTACAGAGCAGAGCAGCGGAGGAGGAG 5520  
Db 5461 GGTGAGACACGACCCATCTGAGAGCTGTGTACAGAGCAGAGCAGCGGAGGAGGAG 5520  
Qy 5521 CAGCGGCGGGGTGGGGGACAGAGCAGCGCGGCTCCCATCTCCCATGACCCACAGCA 5580  
Db 5521 CAGCGGCGGGGTGGGGGACAGAGCAGCGCGGCTCCCATCTCCCATGACCCACAGCA 5580

QY 5581 CTGCCCATCTCCCTCGAAGCCAGAGATGCGCTTCAGACAGAGACCCAGTGTGCTTCAAA 5640  
| | | | |  
Db 5581 CTGCCCATCTCCCTCGAAGCCAGAGATGCGCTTCAGACAGAGACCCAGTGTGCTTCAAA 5640  
| | | | |  
QY 5641 CACAGGCATGAAGGGATATCATCAACGCTGTGAGAGCCAGCAGACCCACGCTCTGAGATC 5700  
| | | | |  
Db 5641 CACAGGCATGAAGGGATATCATCAACGCTGTGAGAGCCAGCAGACCCACGCTCTGAGATC 5700  
| | | | |  
QY 5701 CACCTCACTCTCTCAACCGCTTGGCCAGTGTGCCAATTTCCACCTGTCCACCTGTGCC 5760  
| | | | |  
Db 5701 CACCTCACTCTCTCAACCGCTTGGCCAGTGTGCCAATTTCCACCTGTCCACCTGTGCC 5760  
| | | | |  
QY 5761 ACTGGGCGGACACCTGATGGGGTCTACCTTACCTCATGTAGAGCCGCTTGTGTCGCAAA 5820  
| | | | |  
Db 5761 ACTGGGCGGACACCTGATGGGGTCTACCTTACCTCATGTAGAGCCGCTTGTGTCGCAAA 5820  
| | | | |  
QY 5821 GGAAGCCCCCGGGGTGCGCCCGGCAGAGGGGCCCGAGAGACACCGGGCCATGCTTTCCT 5880  
| | | | |  
Db 5821 GGAAGCCCCCGGGGTGCGCCCGGCAGAGGGGCCCGAGAGACACCGGGCCATGCTTTCCT 5880  
| | | | |  
QY 5881 CGCCAAAGCCCCCGGCTCCGGGCTGAGAGCCGCTCTCTCCCGAGCAAGGGCTTCGGA 5940  
| | | | |  
Db 5881 CGCCAAAGCCCCCGGCTCCGGGCTGAGAGCCGCTCTCTCCCGAGCAAGGGCTTCGGA 5940  
| | | | |  
QY 5941 GCGCCGCGCCCTAGTGTCTCTGTCTGTGGCAAGCCATGCGCCGCAACCCCTTGCGAA 6000  
| | | | |  
Db 5941 GCGCCGCGCCCTAGTGTCTCTGTGTGGCAAGCCATGCGCCGCAACCCCTTGCGAA 6000  
| | | | |  
QY 6001 GAACCTGACACTTCAACAGCGACAGCCCGGACCCGCGGCGCACCTGTGCTCGGAC 6060  
| | | | |  
Db 6001 GAACCTGACACTTCAACAGCGACAGCCCGGACCCGCGGCGCACCTGTGCTCGGAC 6060  
| | | | |  
QY 6061 CCGCGCACCGGAAAAAAGTCTCAAAAGTAAACCTTTTCATCTCAAGAACTGGAACTCGTTC 6120  
| | | | |  
Db 6061 CCGCGCACCGGAAAAAAGTCTCAAAAGTAAACCTTTTCATCTCAAGAACTGGAACTCGTTC 6120  
| | | | |  
QY 6121 TCTGGGTTACACGAGAGAGCTACAGCCCGGAGGGGTGAGGCCGTGACGCCCTGTGAG 6180  
| | | | |  
Db 6121 TCTGGGTTACACGAGAGAGCTACAGCCCGGAGGGGTGAGGCCGTGACGCCCTGTGAG 6180  
| | | | |  
QY 6181 CTTCACCCAGTCTGACCCACAGACAAAGGGGCTCCCAAGCACTTGAAGAGCTCGACAAGAG 6240  
| | | | |  
Db 6181 CTTCACCCAGTCTGACCCACAGACAAAGGGGCTCCCAAGCACTTGAAGAGCTCGACAAGAG 6240  
| | | | |  
QY 6241 CCACTTGAAGGGAGAGCTGCGGCTCAAGCAAGCCAGCCCTGTGAAGCTTGGCGGGAGGC 6300  
| | | | |  
Db 6241 CCACTTGAAGGGAGAGCTGCGGCTCAAGCAAGCCAGCCCTGTGAAGCTTGGCGGGAGGC 6300  
| | | | |  
QY 6301 CGCCCACTCCCAACCTGCGGCGGCTGTGAGAGCCAGCCCTGTGACAGCCCGTGTGCT 6360  
| | | | |  
Db 6301 CGCCCACTCCCAACCTGCGGCGGCTGTGAGAGCCAGCCCTGTGACAGCCCGTGTGCT 6360  
| | | | |  
QY 6361 CCAAGACCGGCCAGGGGTCAAAAGTCAACAGCGGGGTGTCAACCTGTGGCCAGACATCAG 6420  
| | | | |  
Db 6361 CCAAGACCGGCCAGGGGTCAAAAGTCAACAGCGGGGTGTCAACCTGTGGCCAGACATCAG 6420  
| | | | |  
QY 6421 TGAAGTCAATCAACAGAGCTTACACCGGCAACCAACAGCAGCTCAGGCAACCTGTGCC 6480  
| | | | |  
Db 6421 TGAAGTCAATCAACAGAGCTTACACCGGCAACCAACAGCAGCTCAGGCAACCTGTGCC 6480  
| | | | |  
QY 6481 CGCCCCCTCTACTCTCTTCCCTGTGGGCGAGCTGCCCCGCTGTGAGACTTCCGCGCCACAC 6540  
| | | | |  
Db 6481 CGCCCCCTCTACTCTCTTCCCTGTGGGCGAGCTGCCCCGCTGTGAGACTTCCGCGCCACAC 6540  
| | | | |  
QY 6541 CAGTGAAGCTTACTCTCCGCGCCCGAGACATGTGTGCCGCGCTGTGCTCCGCCACAG 6600  
| | | | |  
Db 6541 CAGTGAAGCTTACTCTCCGCGCCCGAGACATGTGTGCCGCGCTGTGCTCCGCCACAG 6600  
| | | | |  
QY 6601 CGAAGGGGCGAAGAGGTCTTCAGAGCCAAACAAAGAGTGTGGTGTGGTGTGGAGAG 6660  
| | | | |  
Db 6601 CGAAGGGGCGAAGAGGTCTTCAGAGCCAAACAAAGAGTGTGGTGTGGTGTGGAGAG 6660  
| | | | |

QY 6661 CGGTATTGAACCTGTGTCTCCCAACCGGAGGCAATGACGAGCCAGGGCACTCCCGAGTGC 6720  
| | | | |  
Db 6661 CGGTATTGAACCTGTGTCTCCCAACCGGAGGCAATGACGAGCCAGGGCACTCCCGAGTGC 6720  
| | | | |  
QY 6721 TGTGTACCGGCTGTGTATCTGATCGGGATGGGAAACAGACGAGCCCAAGATGGGCTTCAA 6780  
| | | | |  
Db 6721 TGTGTACCGGCTGTGTATCTGATCGGGATGGGAAACAGACGAGCCCAAGATGGGCTTCAA 6780  
| | | | |  
QY 6781 GTCTCAGGCAACACAGCAGCCGCGACCTTCTTCAAGACCTCACGAGAGCAATC 6840  
| | | | |  
Db 6781 GTCTCAGGCAACACAGCAGCCGCGACCTTCTTCAAGACCTCACGAGAGCAATC 6840  
| | | | |  
QY 6841 CGCCATGTCTCAAGTCCAAAGAGCAAGATCAACAAAGAGCTGAACCAACCAACCGGAA 6900  
| | | | |  
Db 6841 CGCCATGTCTCAAGTCCAAAGAGCAAGATCAACAAAGAGCTGAACCAACCAACCGGAA 6900  
| | | | |  
QY 6901 TGAAGCTGAATACATATATGACCGACCTGGAGAGAGATCTTCAATATGCGCGCATCAC 6960  
| | | | |  
Db 6901 TGAAGCTGAATACATATATGACCGACCTGGAGAGAGATCTTCAATATGCGCGCATCAC 6960  
| | | | |  
QY 6961 CGGAACAGGCTTTATATGACTTATGAAGCCAGCGGTGTGAGAAACATGCCAGCACCAACT 7020  
| | | | |  
Db 6961 CGGAACAGGCTTTATATGACTTATGAAGCCAGCGGTGTGAGAAACATGCCAGCACCAACT 7020  
| | | | |  
QY 7021 GGGGCTGAGAGCCATTAATTAGAAAGGCACTATGGGTAAATATGACAGTGGGAAGTCT 7080  
| | | | |  
Db 7021 GGGGCTGAGAGCCATTAATTAGAAAGGCACTATGGGTAAATATGACAGTGGGAAGTCT 7080  
| | | | |  
QY 7081 CCGCGCGCTCAAGCGCCATGTCTTTTAAACCTCTGAATGCGCAGTGCAGCGCTGCGTGC 7140  
| | | | |  
Db 7081 CCGCGCGCTCAAGCGCCATGTCTTTTAAACCTCTGAATGCGCAGTGCAGCGCTGCGTGC 7140  
| | | | |  
QY 7141 TATGCCATTAACCGGTGTGAGAGAGAGTGAACCAACATCTACTCTGCGCAGGTGGCGG 7200  
| | | | |  
Db 7141 TATGCCATTAACCGGTGTGAGAGAGAGTGAACCAACATCTACTCTGCGCAGGTGGCGG 7200  
| | | | |  
QY 7201 CCGGAAAGGCGAAAGTCTCTGGGAGAGCCAGCAGCCGAAAGCAATCCCGGCGCCGCGG 7260  
| | | | |  
Db 7201 CCGGAAAGGCGAAAGTCTCTGGGAGAGCCAGCAGCCGAAAGCAATCCCGGCGCCGCGG 7260  
| | | | |  
QY 7261 CCTGGCATCTGGGAGACCGGACCAACCTCTGTCTCTTCAAGTGCATCTCGAGAGAGACTGCAA 7320  
| | | | |  
Db 7261 CCTGGCATCTGGGAGACCGGACCAACCTCTGTCTCTTCAAGTGCATCTCGAGAGAGACTGCAA 7320  
| | | | |  
QY 7321 CCGCGGAGAGCGGCTGTACCAACCGGCTGTGGAGAGACAGGCCCTGTGCGCAGGTTCCAC 7380  
| | | | |  
Db 7321 CCGCGGAGAGCGGCTGTACCAACCGGCTGTGGAGAGACAGGCCCTGTGCGCAGGTTCCAC 7380  
| | | | |  
QY 7381 GCGATTCCCTTCAACCCCTGTATCATGCGGCTGACAGGCGGGGTGTGATGAGCTTCCCAAC 7440  
| | | | |  
Db 7381 GCGATTCCCTTCAACCCCTGTATCATGCGGCTGACAGGCGGGGTGTGATGAGCTTCCCAAC 7440  
| | | | |  
QY 7441 CCCACCGGAGCTCCCGCGGAGGAGCGGCGCCCTGTGCTGCGCCCAACAGCCTGTGAGACGA 7500  
| | | | |  
Db 7441 CCCACCGGAGCTCCCGCGGAGGAGCGGCGCCCTGTGCTGCGCCCAACAGCCTGTGAGACGA 7500  
| | | | |  
QY 7501 GGAAGCCAAAGCACTGTCTGTCTGCAATGAGAGCACTTCCGACAGGAGTGAATCAG 7560  
| | | | |  
Db 7501 GGAAGCCAAAGCACTGTCTGTCTGCAATGAGAGCACTTCCGACAGGAGTGAATCAG 7560  
| | | | |  
QY 7561 AACAAGGCGGG 7620  
| | | | |  
Db 7561 AACAAGGCGGG 7620  
| | | | |  
QY 7621 CGGCGCTGTGAGAGCGGGGCGGCTGTGCGCACTTCCCAACCAAGAGAGCCCTGTGAGTCT 7680  
| | | | |  
Db 7621 CGGCGCTGTGAGAGCGGGGCGGCTGTGCGCACTTCCCAACCAAGAGAGCCCTGTGAGTCT 7680  
| | | | |  
QY 7681 CGGCGCGCTGTGATCTATCTGTCCGTCTGAGAGCGGGGATCTTGTGCTGTCTTAAAGCTT 7740  
| | | | |  
Db 7681 CGGCGCGCTGTGATCTATCTGTCCGTCTGAGAGCGGGGATCTTGTGCTGTCTTAAAGCTT 7740  
| | | | |  
QY 7741 AACTAAGACTCCGCGCCGAGGCTGTGCTGTGACAGACTTACTAGGGGATGTTTACTCTG 7800  
| | | | |







182 CCCGACGCGGAGGCCCTCCCTGCTGTCTGAGTTCACGCCGGGAATGAACGGTCCAG 241  
181 CCCGAGGGGGAGGGCCCTCCCTGCTGTCTGAGTTCAGGCCGGGAATGAACGGTCCAG 240  
242 GAGCTCACTGTGGGCGAGAGTCCCACTCATCTGCGCGAGCTGGGGAAGTCAAGATG 301  
241 GAGCTCACTGTGGGCGAGAGTCCCACTCATCTGCGCGAGCTGGGGAAGTCAAGATG 300  
302 GAGTTCAATTGAAGCAAGCCGCTCGAGCTGCTGCTGACCCCTGCTGCGACG 361  
301 GAGTTCAATTGAAGCAAGCCGCTCGAGCTGCTGCTGACCCCTGCTGCGACG 360  
362 TCACCCCTGTGCGCAACGCGCCAGCTGCGGATCTGAAGACCTCAACAAAGACCGTAGC 421  
361 TCACCCCTGTGCGCAACGCGCCAGCTGCGGATCTGAAGACCTCAACAAAGACCGTAGC 420  
422 CTGACGGGCGAAGCTGGAACGGGTGTCTCCCCCGACGCCCCCGCACTGACCTTGAGCTG 481  
421 CTGACGGGCGAAGCTGGAACGGGTGTCTCCCCCGACGCCCCCGCACTGACCTTGAGCTG 480  
482 GAGCTGTGCGCGCACGGCTGTCCAAAGAGAGCTGATCCAGAACATGACCGGCTGAGAC 541  
481 GAGCTGTGCGCGCACGGCTGTCCAAAGAGAGCTGATCCAGAACATGACCGGCTGAGAC 540  
542 CGAGAGATCACTATGTTAGAGCAGACAGATCTTAAGCTGAAGAAAGACGACCAACGCTG 601  
541 CGAGAGATCACTATGTTAGAGCAGACAGATCTTAAGCTGAAGAAAGACGACCAACGCTG 600  
602 GAGGAGAGAGCTGCCAAGCCCGCGAGCTGAGAAAGCCGTGTCAACGCGCCCATGAG 661  
601 GAGGAGAGAGCTGCCAAGCCCGCGAGCTGAGAAAGCCGTGTCAACGCGCCCATGAG 660  
662 TCGAAGACCCGACGCTGTGAGATCATATGAGAGAAACCGGAAGAGGCTGAAGCT 721  
661 TCGAAGACCCGACGCTGTGAGATCATATGAGAGAAACCGGAAGAGGCTGAAGCT 720  
722 GCACATCGAATTTCTGGAAGGCTGCGGCCCCAGGTGAGCTGCGCTGTACCAACGAGCC 781  
721 GCACATCGAATTTCTGGAAGGCTGCGGCCCCAGGTGAGCTGCGCTGTACCAACGAGCC 780  
782 TCCGACACCCGCGAGTATCATGGAACATCAAAATTAACAGGCGATGCGGAAGAGCTA 841  
781 TCCGACACCCGCGAGTATCATGGAACATCAAAATTAACAGGCGATGCGGAAGAGCTA 840  
842 ATCTTGAATTTCAAGAGGAGGAAATCAAGCTCGAAGCAATGAGAGAGAGTTCTGCGAG 901  
841 ATCTTGAATTTCAAGAGGAGGAAATCAAGCTCGAAGCAATGAGAGAGAGTTCTGCGAG 900  
902 CGCTATGACCAAGCTCATGAGAGGCTTGAAGAAAGAGTGAAGCGCATCGAAGAACCCG 961  
901 CGCTATGACCAAGCTCATGAGAGGCTTGAAGAAAGAGTGAAGCGCATCGAAGAACCCG 960  
962 CGCGGCGGCGCAAGAGAGCAAGGTGCGGAGTACTAGAAAGCAATTCCTTGAGATC 1021  
961 CGCGGCGGCGCAAGAGAGCAAGGTGCGGAGTACTAGAAAGCAATTCCTTGAGATC 1020  
1022 CGGAACAGCGCGAGCTGACGAGAGCGCATGACAGAGGTTGGCCAGCGCGGCGAGTGG 1081  
1021 CGGAACAGCGCGAGCTGACGAGAGCGCATGACAGAGGTTGGCCAGCGCGGCGAGTGG 1080  
1082 CTGTCAATGTGCGCGCGCGCGAGCAGAGCAAGAGTGTCAAGAGATCATGATGCGCTTCA 1141  
1081 CTGTCAATGTGCGCGCGCGCGAGCAGAGCAAGAGTGTCAAGAGATCATGATGCGCTTCA 1140  
1142 GAGCAGAGAACTGTGAAGAGCAGATGCGCAGCTGCGCTGTATCCCGCCCATGCTGATC 1201  
1141 GAGCAGAGAACTGTGAAGAGCAGATGCGCAGCTGCGCTGTATCCCGCCCATGCTGATC 1200  
1202 GAGCTGACCAAGAGAGGCAATCAAGTTCAATCAATGAACGGGCTTATGCGCGAGCCCATG 1261  
1201 GAGCTGACCAAGAGAGGCAATCAAGTTCAATCAATGAACGGGCTTATGCGCGAGCCCATG 1260  
1262 AAGGTATCAAAAGACGCGCAGGTCAATGAACATGTGAGTGAAGAGAGAGACCTTTC 1321

1261 AAGGTATCAAAAGACGCGCAGGTCAATGAACATGTGAGTGAAGAGAGAGACCTTTC 1320  
1322 CGGGAAGAGTTCAATGAGACATCCCAAGAACTTTGGCTGATGCAATCATCTTGAGAGG 1381  
1321 CGGGAAGAGTTCAATGAGACATCCCAAGAACTTTGGCTGATGCAATCATCTTGAGAGG 1380  
1382 AAGCAGTGTGATGATGCGCTCTCTATTAATCACTGAATGAAGAAATGAATATTAAG 1441  
1381 AAGCAGTGTGATGATGCGCTCTCTATTAATCACTGAATGAAGAAATGAATATTAAG 1440  
1442 AGCTGTGATGATGAGAGCTATGCGCGCGCGGCAAGAGC---CAGCAGCAACACAGCAG 1498  
1441 AGCTGTGATGATGAGAGCTATGCGCGCGCGGCAAGAGCAGCAGCAGCAACACAGCAG 1500  
1499 CAGCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1558  
1501 CAGCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
1559 AAGATGAG 1618  
1561 AAGATGAG 1620  
1619 AACGACAG 1678  
1621 AACGACAG 1680  
1679 GAG 1738  
1681 GAG 1740  
1739 GAGCGCATCAACCGCTCAATGAGCTAATGAGGCGCAACAGCAGAGAGAGAGAGAGAG 1798  
1741 GAGCGCATCAACCGCTCAATGAGCTAATGAGGCGCAACAGCAGAGAGAGAGAGAGAG 1800  
1799 CAGAGCGCGAGAGCTGCGCTTCATGAGAGCTGATGAGAGAGAGAGAGAGAGAGAGAG 1858  
1801 CAGAGCGCGAGAGCTGCGCTTCATGAGAGCTGATGAGAGAGAGAGAGAGAGAGAGAG 1860  
1859 ATGGAACAGCCCAAGAAAG 1918  
1861 ATGGAACAGCCCAAGAAAG 1920  
1919 ATGAGTGGCTCCAAAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1978  
1921 ATGAGTGGCTCCAAAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980  
1979 CAGAACCTGATGAGATCTTTCAGCAGCACAAGCTGAAGATGAGAGAGAGAGAGAGAGAG 2038  
1981 CAGAACCTGATGAGATCTTTCAGCAGCACAAGCTGAAGATGAGAGAGAGAGAGAGAGAG 2040  
2039 CGAGAGAAAG 2098  
2041 CGAGAGAAAG 2100  
2099 GTGAGAGATGAGAGAGATGAG 2158  
2101 GTGAGAGATGAGAGAGATGAG 2160  
2159 GAGGCTGAAGCTTTCATGAG 2218  
2161 GAGGCTGAAGCTTTCATGAG 2220  
2219 GCGACTGTCAACACAGCTCAGACACCGAGAGAGATCCCTCTCTCTCACTGAGAGCGCGC 2278  
2221 GCGACTGTCAACACAGCTCAGACACCGAGAGAGATCCCTCTCTCTCACTGAGAGCGCGC 2280  
2279 AAGGACAGAGGAG 2338  
2281 AAGGACAGAGGAG 2340  
2339 CAGGCGCAACCAACCCACACGAGAGAGATCCCGGCGCCCAATGAGCCCAACCCGCGCC 2398

D 2341 CCAAGCCACCCACCCACGAGGACATCCCGGGCCCCCATTTAGCCACCCCGGCC 2400  
Q 2399 TCTGAAGCCACCGAGCCCTTAGCGCCCCACGACACCCCAATGCGCTCTGCACTCTT 2458  
D 2401 TCTGAAGCCACCGAGCCCTTAGCGCCCCACGAGCACCCCAATGCGCTCTGCACTCTT 2460  
Q 2459 CCTGTGTCTCCCAAGAGAGAAAGAGAGAGACCGACAGCGCCCCAGTGAAGAG 2518  
D 2461 CTTGTGTCTCCCAAGAGAGAGAGAGAGACCGACAGCGCCCCAGTGAAGAG 2520  
Q 2519 GGGGAGAGACAAAGCCCCCGCTGAGAGCTGGGCACTGGACAAGAGAGCGGAG 2578  
D 2521 GGGGAGAGACAAAGCCCCCGCTGAGAGCTGGGCACTGGACAAGAGAGCGGAG 2580  
Q 2579 GAGCCCTCAAGAGAGAGTGCACGAGAGAACGAGAGAGGGCGCGGCAAGGGCAAGAC 2638  
D 2581 GAGCCCTCAAGAGAGAGTGCACGAGAGAACGAGAGAGGGCGCGGCAAGGGCAAGAC 2640  
Q 2639 GCGAGAGCGCTGAGAGCCAGCGCCGAGGGGGCGCTCAAGGCAAGAGAGAGAGGGG 2698  
D 2641 GCGAGAGCGCTGAGAGCCAGCGCCGAGGGGGCGCTCAAGGCAAGAGAGAGGGG 2700  
Q 2699 AGCGGACGGGGCCACCATGCGCAAGAGCTGGGGCGCCCCAGAGACAGCATCCAGTGT 2758  
D 2701 AGCGGACGGGGCCACCATGCGCAAGAGCTGGGGCGCCCCAGAGACAGCATCCAGTGT 2760  
Q 2759 ACCTGACGTGCAGAGAGGTGATGAGCGGAGGGCGGCGAGCAAGAAACGGCTGTCTC 2818  
D 2761 ACCTGACGTGCAGAGAGGTGATGAGCGGAGGGCGGCGAGCAAGAACGGGTGTCTC 2820  
Q 2819 CCAAGGCCACGCTCTCTACCCCCGACTGGCGACCCCCGGGCGCAATGCTCACTCCAGAG 2878  
D 2821 CCAAGGCCACGCTCTCTACCCCCGACTGGCGACCCCCGGGCGCAATGCTCACTCCAGAG 2880  
Q 2879 CCACTGACCTGAAGAGCTGAAGCAGGAGGGGCTGCATCCCCCGCATTCAGAGTCAAC 2938  
D 2881 CCACTGACCTGAAGAGCTGAAGCAGGAGGGGCTGCATCCCCCGCATTCAGAGTCAAC 2940  
Q 2939 AAAGTCAATGAGCCCCCGGGAGAGACGAGAGCTCCCAACCAAGCCAGCTCCCGACCCCA 2998  
D 2941 AAAGTCAATGAGCCCCCGGGAGAGACGAGAGCTCCCAACCAAGCCAGCTCCCGACCCCA 3000  
Q 2999 CCGGCAACGGCAAAACCTGCAAGCGGAGAGACGCCCCCTCAGAGACTGGCAGACGCC 3058  
D 3001 CCGGCAACGGCAAAACCTGCAAGCGGAGAGACGCCCCCTCAGCAACCTGGCAGACGCC 3060  
Q 3059 CGGGGCAAGAGAGAGCCCGGACCCCCCGGACCAAGAGAGCCCTTGGCAACCGAGAGGC 3118  
D 3061 CGGGGCAAGAGAGAGCCCGGACCCCCCGGACCAAGAGAGCCCTTGGCAACCGAGAGGC 3120  
Q 3119 CAGAAAGCTGAGGAGACCCCTTGTCTGAGCTTCCGAGCTGCCCCCTGAGCCCCC 3178  
D 3121 CAGAAAGCTGAGGAGACCCCTTGTCTGAGCTTCCGAGCTGCCCCCTGAGCCCCC 3180  
Q 3179 CGTGAAGTGAATAAAGCCTCCCGCATGCCCCGAGACCCCTCAGCCTTCTCTAACCTCA 3238  
D 3181 CGTGAAGTGAATAAAGCCTCCCGCATGCCCCGAGACCCCTCAGCCTTCTCTAACCTCA 3240  
Q 3239 CTTGATCAACCACTGAGCCCTTGGGCGTCAATGACATGCCCCGGGCGCTTGGCGGGCCA 3298  
D 3241 CTTGATCAACCACTGAGCCCTTGGGCGTCAATGACATGCCCCGGGCGCTTGGCGGGCCA 3300  
Q 3299 CCAACCATCTCCAAACCGGCTCCCTCATCTCTCTGCAAGACCCCAAGCGTCTCTGAG 3358  
D 3301 CCAACCATCTCCAAACCGGCTCCCTCATCTCTCTGCAAGACCCCAAGCGTCTCTGAG 3360  
Q 3359 AGGCAAAATAGGTGCCATCTCCCAAGAAATGTGGTCAAGCTCAAGTCCGTAATCAAG 3418  
D 3361 AGGCAAAATAGGTGCCATCTCCCAAGAAATGTGGTCAAGCTCAAGTCCGTAATCAAG 3420  
Q 3419 CATGCAAGGCGCGGTGGGCGCTGTCAACATGGGGCTGGCCCTGCAATGAGCCCAAA 3478  
D 3421 CATGCAAGGCGCGGTGGGCGCTGTCAACATGGGGCTGGCCCTGCAATGAGCCCAAA 3480

Q 3479 AAGCTTGGACCTTTAGCGAGTGAAGAGAGAGAGCTGTCCCCACGAGGCCAGGCTGG 3538  
D 3481 AAGCTTGGACCTTTAGCGAGTGAAGAGAGAGAGAGCTGTCCCCACGAGGCCAGGCTGG 3540  
Q 3539 CCAACCGAGAGCTTGGGGTGTCCCAAGCCCAAGAGAGGCTCGTGTGAAGAGAGAGCT 3598  
D 3541 CCAACCGAGAGCTTGGGGTGTCCCAAGCCCAAGAGAGGCTCGTGTGAAGAGAGAGCT 3600  
Q 3599 CTGGCTGATTCGGGGGGAAGCATCAACAAAGGCAATTCACAGACAAGGTTGCCCTGG 3658  
D 3601 CTGGCTGATTCGGGGGGAAGCATCAACAAAGGCAATTCACAGACAAGGTTGCCCTGG 3660  
Q 3659 GACAGCGCATCAATACCGGCGCTCATCAACCAAGGACCGGCACTGACGCTCTGTAC 3718  
D 3661 GACAGCGCATCAATACCGGCGCTCATCAACCAAGGACCGGCACTGACGCTCTGTAC 3720  
Q 3719 AAGGCAACATCAACAGATCATCGGCGAGAGACGCCGAGTCCGTTGAGCTGCGCGGG 3778  
D 3721 AAGGCAACATCAACAGATCATCGGCGAGAGACGCCGAGTCCGTTGAGCTGCGCGGG 3780  
Q 3779 GAGGACAGCTGGCCCAAGGGCCAGTCAATCAAGAAAGCAAGAGGGCCAGTCTGTCC 3838  
D 3781 GAGGACAGCTGGCCCAAGGGCCAGTCAATCAAGAAAGCAAGAGGGCCAGTCTGTCC 3840  
Q 3839 TATGAGGTGGCATGTCTGTGACCCAGTGTCCAAAGAGAGACGGCAGAGACGCTCAAG 3898  
D 3841 TATGAGGTGGCATGTCTGTGACCCAGTGTCCAAAGAGAGACGGCAGAGACGCTCAAG 3900  
Q 3899 CCCCCCATGAGACGGCGCCCCCAAGGCGACCTTACATATATATGAAGGCGCGTGGG 3958  
D 3901 CCCCCCATGAGACGGCGCCCCCAAGGCGACCTTATATATATATGAAGGCGCGTGGG 3960  
Q 3959 AAGGCATCTCCTGACGAGCATCGAAGGTCTCATGTGGGCGGTCATCCGCGGAGGCA 4018  
D 3961 AAGGCATCTCCTGACGAGCATCGAAGGTCTCATGTGGGCGGTCATCCGCGGAGGCA 4020  
Q 4019 CACAGCCCCCAACCACTTAAAGAGCAGACCAATCCGCGGTTCATCAACAAAGGATC 4078  
D 4021 CACAGCCCCCAACCACTTAAAGAGCAGACCAATCCGCGGTTCATCAACAAAGGATC 4080  
Q 4079 CCTCGGTCTTAAGTGAAGCACAAGAGACCTACCTGCGTGGGAGGCCAAGTCTTAAAG 4138  
D 4081 CCTCGGTCTTAAGTGAAGCACAAGAGACCTACCTGCGTGGGAGGCCAAGTCTTAAAG 4140  
Q 4139 CGGAGGAGCAGGCTCGGCCCCACCGGCTCAAGGAGCTGAACGAGGCGCTTACAAGAG 4198  
D 4141 CGGAGGAGCAGGCTCGGCCCCACCGGCTCAAGGAGCTGAACGAGGCGCTTACAAGAG 4200  
Q 4199 CAGGCTTGGGCCCCCTGAAGCTGAAGCGGCGCATGAAGGCGCTGTGGCCACGTTGAAG 4258  
D 4201 CAGGCTTGGGCCCCCTGAAGCTGAAGCGGCGCATGAAGGCGCTGTGGCCACGTTGAAG 4260  
Q 4259 GAGCGGGGCGCTCATCATGAAGTCCCGCGGAGAGACTGCGGCAACGCCGAGCTG 4318  
D 4261 GAGCGGGGCGCTCATCATGAAGTCCCGCGGAGAGACTGCGGCAACGCCGAGCTG 4320  
Q 4319 CCGCTGAGCCCGGCGGCTCAAGAGAGGCTCAATCAAGCAGGAGACCCCGTCAAGTAC 4378  
D 4321 CCGCTGAGCCCGGCGGCTCAAGAGAGGCTCAATCAAGCAGGAGACCCCGTCAAGTAC 4380  
Q 4379 GACACCGGCGGCTCAACCATGAGCTCAAAAAGACAGAGTACGTTCCCTCATCGGACG 4438  
D 4381 GACACCGGCGGCTCAACCATGAGCTCAAAAAGACAGAGTACGTTCCCTCATCGGACG 4440  
Q 4439 CCGGCGCGAGGTTCCACCCGCTGACCCGCTGATGTGATGAGCGGACCGCGGCACTG 4498  
D 4441 CCGGCGCGAGGTTCCACCCGCTGACCCGCTGATGTGATGAGCGGACCGCGGCACTG 4500  
Q 4499 GAAAGTGTCTGTACGAGAGAGCTGAAGAGCGGCGCAGAGGACCGGACGAGCTGGGG 4558  
D 4501 GAAAGTGTCTGTACGAGAGAGCTGAAGAGCGGCGCAGAGGACCGGACGAGCTGGGG 4560

4559 GGCCTCATTTGCGCGCGCGCCCGCTCATTTGCTGAGCTGTGAAGCCGCGCAGAGC 4618  
4561 GGGTTCATTGCGCGCGCGCCCGCTCATTTGCTGAGCTGTGAAGCCGCGCAGAGC 4620  
4619 CCCCTGACCTATGAGACCAACGCGGACACCTTTGCGGCGCACTCCACGAGGTTGCCC 4678  
4621 CCCCTAACCTATGAGACCAACGCGGACACCTTTGCGGCGCACTCCACGAGGTTGCCC 4680  
4679 GTGACCATGGGGAGGCCACCGCGCGCTGAGAGGCGGCGCTTTGCTCAACAGGCA 4738  
4681 GTGACCAACGGGGAGGCCACCGCGCGCTGAGAGGCGGCGCTTTGCTCAACAGGCA 4740  
4739 TCCCAAGAACGAAAGTGAAGTGAAGCTGAGAGCTGAGATCGCAAGTCCCGCACAAGACC 4798  
4741 TCCCAAGAACGAAAGTGAAGTGAAGCTGAGAGCTGAGATCGCAAGTCCCGCACAAGACC 4800  
4799 GTGCGCGAGACCAACCAACCCCATCTGCGCTTATGAGCACTTGCTTGGGGCTGAGT 4858  
4801 GTGCGCGAGACCAACCAACCCCATCTGCGCTTATGAGCACTTGCTTGGGGCTGAGT 4860  
4859 GGGGTGAGACCTGTATGAGAGGCAATCCCGCTGGGCTTGAACCCACCTGCATTAACCCGC 4918  
4861 GGGGTGAGACCTGTATGAGAGGCAATCCCGCTGGGCTTGAACCCACCTGCATTAACCCGC 4920  
4919 GGCATCCTCTGTGAGACGAGCGCTGCTACTTACTGCGCGCAGACCTGCGCCCAACCC 4978  
4921 GGCATCCTCTGTGAGACGAGCGCTGCTACTTACTGCGCGCAGACCTGCGCCCAACCC 4980  
4979 ACCTACCCGCACTGTATCCCAACCTTACCTTATCCGCGGCTTACCCGACAGCGCGGCTG 5038  
4981 ACCTACCCGCACTGTATCCCAACCTTACCTTATCCGCGGCTTACCCGACAGCGCGGCTG 5040  
5039 GAGAACCGGAGACCAATCATCATATGATATCATCTCGAGAGAGATGACCAACAACG 5098  
5041 GAGAACCGGAGACCAATCATCATATGATATCATCTCGAGAGAGATGACCAACAACG 5100  
5099 GCCACCGCACTGAGCGAGCTGATATGCTGAGGGGCTCTGCGCCCGCGAGTCTCG 5158  
5101 GCCACCGCACTGAGCGAGCTGATATGCTGAGGGGCTCTGCGCCCGCGAGTCTCG 5160  
5159 CTGGCACTCAACTGACCTGCGGGTCCCGAGGACATGACCTGTCTCCAAAGTGCACAC 5218  
5161 CTGGCACTCAACTGACCTGCGGGTCCCGAGGACATGACCTGTCTCCAAAGTGCACAC 5220  
5219 CTGCTGTGTCTGTGCTGCTGCGCGACCAAGAGCAACCCGACCTGATGACCGCTTGGC 5278  
5221 CTGCTGTGTCTGTGCTGCTGCGCGACCAAGAGCAACCCGACCTGATGACCGCTTGGC 5280  
5279 TACCTCCCAACGCGCGCCGACCTTCAAGAGCGGACCAAGAGCTTCCCACTTCCCA 5338  
5281 TACCTCCCAACGCGCGCCGACCTTCAAGAGCGGACCAAGAGCTTCCCACTTCCCA 5340  
5339 GAGAGTCAACACACTTGAACAAACCAACCAACGACCTGCTCTGTCGAGCGGAGCGAGAC 5398  
5341 GAGAGTCAACACACTTGAACAAACCAACCAACGACCTGCTCTGTCGAGCGGAGCGAGAC 5400  
5399 CGGGATTCGAGAGCGGAGCGGGATTCGGAGCGGGAGAAAGTCTACTCTCACTCAACG 5458  
5401 CGGGATTCGAGAGCGGAGCGGGATTCGGAGCGGGAGAAAGTCTACTCTCACTCAACG 5460  
5459 ACGGTGAGACGACCACTTGAAGCTGTGAGAGCTGTGAGAGAGAGAGAGAGAGAGAG 5518  
5461 ACGGTGAGACGACCACTTGAAGCTGTGAGAGCTGTGAGAGAGAGAGAGAGAGAGAG 5520  
5519 AGCAGCGCGGGGTGGGGAGAGAGAGCGCGCGCTTCCACTTCCCAATGCGCACG 5578  
5521 AGCAGCGCGGGGTGGGGAGAGAGAGCGCGCGCTTCCACTTCCCAATGCGCACG 5580  
5579 CACTCGCCCATCTTCCCTCGGAGCCGAGATGCTTCAAGAGAGAGAGAGAGAGAGAGAG 5638  
5581 CACTCGCCCATCTTCCCTCGGAGCCGAGATGCTTCAAGAGAGAGAGAGAGAGAGAGAG 5640  
5639 AACACAGCATGAAGGTATCATCAACGCTGTGAGAGCCAGAACCCACGCTCTGAGG 5698

5641 AACACAGCATGAAGGTATCATCAACGCTGTGAGAGCCAGACCGCTCATGAGG 5700  
5699 TCCACTCAACCTCTCAACCCGCTTGGCGGAGAGCTGAGCAATTCACCTGCAACCACTG 5758  
5701 TCCACTCAACCTCTCAACCCGCTTGGCGGAGAGCTGAGCAATTCACCTGCAACCACTG 5760  
5759 CCACTGGGCGGCAACCTTCAATGAGGATCTACCTTACCTTCAATGAGAGCGCTTGTGCCC 5818  
5761 CCACTGGGCGGCAACCTTCAATGAGGATCTACCTTACCTTCAATGAGAGCGCTTGTGCCC 5820  
5819 AAGAGAGCCCGCGGGTCCCGCGGAGAGAGCGGCGCGGAGAGAGAGAGAGAGAGAGAG 5878  
5821 AAGAGAGCCCGCGGGTCCCGCGGAGAGAGCGGCGCGGAGAGAGAGAGAGAGAGAGAG 5880  
5879 CTGCGCAACCCCAACCCGCTTCCGGGCTGAGAGCGGCGCTTCCCGCGGAGAGAGAGAG 5938  
5881 CTGCGCAACCCCAACCCGCTTCCGGGCTGAGAGCGGCGCTTCCCGCGGAGAGAGAGAG 5940  
5939 GAGCCCGGCGCTTATGCTCTGCTCTGAGGCAAGCAACATCGCCGCAACCCCTGCG 5998  
5941 GAGCCCGGCGCTTATGCTCTGCTCTGAGGCAAGCAACATCGCCGCAACCCCTGCG 6000  
5999 AAGAACCTGCACTTCAACAGCGAGCGGAGCGGCGCGGAGCGGAGCGGAGCGGAGCGG 6058  
6001 AAGAACCTGCACTTCAACAGCGAGCGGAGCGGCGCGGAGCGGAGCGGAGCGGAGCGG 6060  
6059 GACCGGCAACGGGAGAAAGATCAAAATGAACCTTTTCCATCCAGAGAACTGGAATCCGT 6118  
6061 GACCGGCAACGGGAGAAAGATCAAAATGAACCTTTTCCATCCAGAGAACTGGAATCCGT 6120  
6119 TCTCTGGGTTACCAAGGAGAGAGCTTCAAGGCGCGGAGAGAGAGAGAGAGAGAGAGAG 6178  
6121 TCTCTGGGTTACCAAGGAGAGAGCTTCAAGGCGCGGAGAGAGAGAGAGAGAGAGAGAG 6180  
6179 AGCTCAACCAAGTGAACCAACGACAAAGGAGCTTCCCAAGCACTGGAAGCTTGAACAG 6238  
6181 AGCTCAACCAAGTGAACCAACGACAAAGGAGCTTCCCAAGCACTGGAAGCTTGAACAG 6240  
6239 AGCCACTGAGAGGGAGAGCTGCGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6298  
6241 AGCCACTGAGAGGGAGAGCTGCGGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6300  
6299 GCCGCCAACCTCCCAACCTGCGGCGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 6358  
6301 GCCGCCAACCTCCCAACCTGCGGCGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 6360  
6359 CTTCAGACCGCCCAAGGGTCAAAAGTCAACAGCGGAGTCACTTGGCTCAGACATC 6418  
6361 CTTCAGACCGCCCAAGGGTCAAAAGTCAACAGCGGAGTCACTTGGCTCAGACATC 6420  
6419 AGTGAAGTCAATCAACAGAGCTTACCGCGGACCAACCAAGAGAGAGAGAGAGAGAGAG 6478  
6421 AGTGAAGTCAATCAACAGAGCTTACCGCGGACCAACCAAGAGAGAGAGAGAGAGAGAG 6480  
6479 CCGGCGCCCTCTACTCTTCTTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6538  
6481 CCGGCGCCCTCTACTCTTCTTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6540  
6539 CCGAGTGAACCTTACTCTTCTTCTTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6598  
6541 CCGAGTGAACCTTACTCTTCTTCTTCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6600  
6599 AGCGAAGGGGAGAGAGAGTCTTCAAGAGCAAAACAAAGAGAGAGAGAGAGAGAGAGAG 6658  
6601 AGCGAAGGGGAGAGAGAGTCTTCAAGAGCAAAACAAAGAGAGAGAGAGAGAGAGAGAG 6660  
6659 GACGATTTGAACCTGATGCTTCCCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6718  
6661 GACGATTTGAACCTGATGCTTCCCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6720  
6719 GCTGTATACCGGCTGTGTATCCGGATGAGGAAACAGAGAGCCAGAGAGAGAGAGAGAG 6778

Db 6721 GCTGTGTACCCGCTGCTGTATCCGGGATGGGGAACAGACGAGCCACGACGATGGGCTCC 6780  
QY AAGTCTCCAGGCAACCCAGCCAGCCAGGCTTCTTCAGCAAGTGAACGAGAGCAAC 6838  
Db 6781 AAGTCTCCAGGCAACCCAGCCAGCCAGGCTTCTTCAGCAAGTGAACGAGAGCAAC 6840  
QY TCCGCTCATGTCAAGTCCAAAGAGCAAGATCAACAAAGAGCTGAAACCCCAACCCG 6898  
Db 6841 TCCGCTCATGTCAAGTCCAAAGAGCAAGATCAACAAAGAGCTGAAACCCCAACCCG 6900  
QY AATGAGCCCTGAATTAATATCAAGCCAGCTGGAGACGAGATCTTCAATATGCCCCATC 6958  
Db 6901 AATGAGCCCTGAATTAATATCAAGCCAGCTGGAGACGAGATCTTCAATATGCCCCATC 6960  
QY ACCGGAACAGGCTTATGACTATAGAACCCAGCGGTGACAGAAACATGSCAGACCAAC 7018  
Db 6961 ACCGGAACAGGCTTATGACTATAGAACCCAGCGGTGACAGAAACATGSCAGACCAAC 7020  
QY ATGGGGCTGGAGGCAATTAAGAAAGGCACTCATGGGTAATATGACCAAGTGGGAAAG 7078  
Db 7021 ATGGGGCTGGAGGCAATTAAGAAAGGCACTCATGGGTAATATGACCAAGTGGGAAAG 7080  
QY TCCCGCCGCTCAGGCGCAATGCTTTTAACTCTTGAAATGCCAGTGGCCAGCTGCGCT 7138  
Db 7081 TCCCGCCGCTCAGGCGCAATGCTTTTAACTCTTGAAATGCCAGTGGCCAGCTGCGCT 7140  
QY GCTATGACCATAACCGCTGCTGACGAGACGAGATGACACACACTCACTCGCCAGGTGC 7198  
Db 7141 GCTATGACCATAACCGCTGCTGACGAGACGAGATGACACACACTCACTCGCCAGGTGC 7200  
QY GCGGGGAAAGGCAAGGTCTCTGGCAGACCAGCAGCCGAAAGCAAGTCCCGGCCCCG 7258  
Db 7201 GCGGGGAAAGGCAAGGTCTCTGGCAGACCAGCAGCCGAAAGCAAGTCCCGGCCCCG 7260  
QY GGCCTGCGATCTGGGGACCCGGACCCCTGTGTCTCTCACTGACGACCTCGGAGGGAATGC 7318  
Db 7261 GGCCTGCGATCTGGGGACCCGGACCCCTGTGTCTCTCACTGACGACCTCGGAGGGAATGC 7320  
QY AACCCGCGGACCGCTCAACACCGCGTGTGGAGAGACAGGCCCTCGTCCGACGATTC 7378  
Db 7321 AACCCGCGGACCGCTCAACACCGCGTGTGGAGAGACAGGCCCTCGTCCGACGATTC 7380  
QY ACGCATTCCTCTACCAACCCCTGTATCATGCGGCTGACAGCGGGGTGATGCTCCCA 7438  
Db 7381 ACGCATTCCTCTCTACCAACCCCTGTATCATGCGGCTGACAGCGGGGTGATGCTCCCA 7440  
QY CCCCCACCGGGCTCCCGCGGACCGGGCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 7498  
Db 7441 CCCCCACCGGGCTCCCGCGGACCGGGCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTG 7500  
QY GAGAGCCCAAGCCACTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7558  
Db 7501 GAGAGCCCAAGCCACTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7560  
QY AGAACAGGCGCGGGCGGGCGGGCGGTBUERSPERLUGTCAAGTCCCAAGCAAGCAGG 7618  
Db 7561 AGAACAGGCGCGGGCGGGCGGGCGGGCGG-----TGTCAAGTCCCAAGCAAGCAGG 7610  
QY AACGCGCTCTGACGAGCGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7678  
Db 7611 AACGCGCTCTGACGAGCGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7670  
QY TCCGCTGCGCTCATCATCTGTCCGTCAGAGACCGGACATCTTGTGCTTAAAGCC 7738  
Db 7671 TCCGCTGCGCTCATCATCTGTCCGTCAGAGACCGGACATCTTGTGCTTAAAGCC 7730  
QY TTAACCTAAGACTCCGCGCCCGGCTGACCTTGTGCAACCTTAACCTCAAGGAGATGTTTACC 7798  
Db 7731 TTAACCTAAGACTCCGCGCCCGGCTGACCTTGTGCAACCTTAACCTCAAGGAGATGTTTACC 7790  
QY TGTGTCTCGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7858  
Db 7791 TGTGTCTCGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7850

QY 7859 ACACACAGGCGGCGACAGGCGGCGACAGGAGCCCAAAACAGAGATGACCGACCTCCAGCGC 7918  
Db 7861 ACACACAGGCGGCGACAGGCGGCGACAGGAGCCCAAAACAGAGATGACCGACCTCCAGCGC 7910  
QY ACTGCTCTCCCGAAATGCAATTTGGAACCAAAAGTCTTAACTGAGCTTGCAGCCCGCGC 7978  
Db 7911 ACTGCTCTCCCGAAATGCAATTTGGAACCAAAAGTCTTAACTGAGCTTGCAGCCCGCGC 7970  
QY CTCCCTCGGCTCCCATCCCGCTTAGCGCTGACAGATGGAAGCAGAGCCCTGTCCAGC 8038  
Db 7971 CTCCCTCGGCTCCCATCCCGCTTAGCGCTTGAAGCAAGATGGAAGCAGAGCCCTGTCCAGC 8030  
QY CCCAGTGCCTGCTTCCGCTTCCCAAGACTGCTCCCAAGCAAGATGCTGGAACCC 8098  
Db 8031 CCCAGTGCCTGCTTCCGCTTCCCAAGACTGCTCCCAAGCAAGATGCTGGAACCC 8090  
QY AAGTCAAGCCAGGTGGCGGACAAAGGCGCAGGTGCGGCTGGGGGGAACGATGCTCC 8158  
Db 8091 AAGTCAAGCCAGGTGGCGGACAAAGGCGCAGGTGCGGCTGGGGGGAACGATGCTCC 8150  
QY GAGGACTGGAATGTTTTTTCACATGCTTGCAGAGGCGGTGGGGAAGGAGAGATG 8218  
Db 8151 GAGGACTGGAATGTTTTTTCACATGCTTGCAGAGGCGGTGGGGAAGGAGAGATG 8210  
QY TAAATGATGTGTGTTTACAGGGTATATTTTGTATACCTTCAATGAATTAATTCAGATG 8278  
Db 8211 TAAATGATGTGTGTTTACAGGGTATATTTTGTATACCTTCAATGAATTAATTCAGATG 8270  
QY TTTTACGAGAGAGAGACTTACCCAGATTTACTGCTGCTGCTTTTGTATCTTCTTAC 8338  
Db 8271 TTTTACGAGAGAGAGACTTACCCAGATTTACTGCTGCTGCTTTTGTATCTTCTTAC 8330  
QY GGTCAAGAGGCGTGTGAGGCGGACAGTGGGTGACCCCATCACTCGAGAGCAAGGGG 8398  
Db 8331 GGTCAAGAGGCGTGTGAGGCGGACAGTGGGTGACCCCATCACTCGAGAGCAAGGGG 8390  
QY GCGGGGACTGCTCG-TCAAGCCCGCTGTGTCTTCCCTCCCTTCCCTTGGGAGAGAT 8457  
Db 8391 GCGGGGACTGCTCGTCAAGCCCGCTGTGTCTTCCCTCCCTTCCCTTGGGAGAGAT 8450  
QY GAATTCGATGCGTATTTCTGTGGCCGCAATTTGCGCAGGGGTGTGTATTTCTGTCAATTAAC 8517  
Db 8451 GAATTCGATGCGTATTTCTGTGGCCGCAATTTGCGCAGGGGTGTGTATTTCTGTCAATTAAC 8510  
QY ACAGTGTCTCAATTAATAAGCGAATTAATCTCCAAAAAAGGAGAGAGAGAGAGAGAGAG 8571  
Db 8511 ACAGTGTCTCAATTAATAAGCGAATTAATCTCCAAAAAAGGAGAGAGAGAGAGAGAGAG 8564

RESULT 4  
ACN44283  
ID ACN44283 standard; cDNA, 8667 BP.  
XX ACN44283;  
XX  
AC ACN44283;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human mRNA sequence hCT16428.  
XX  
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN W02003073826-A2.  
XX  
XX 12-SEP-2003.  
PD  
XX 28-FEB-2003; 2003WO-US006235.  
PF  
XX 01-MAR-2002; 2002US-00087192.  
PR  
XX  
XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;  
XX WPI; 2003-328604/31.  
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma  
XX comprises a nucleotide sequence.  
XX  
XX Claim 1; SEQ ID NO 653; Opp; English.  
XX  
XX The present invention relates to novel DNA and protein sequences which  
XX are associated with carcinomas. The sequences are useful for: (i) for  
XX screening drug candidates; (ii) for screening of bioactive agent capable  
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of  
XX a bioactive agent capable of modulating the activity of CAP; (iv) for  
XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing  
XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating  
XX carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;  
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for  
XX determining Carcinoma Associated (CA) gene copy number. In addition, the  
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of  
XX carcinoma including lymphoma. The present sequence is one such CA gene  
XX sequence. Note: This patent is an equivalent to basic patent  
XX US2002182586A1, for which no sequence data was published  
SQ Sequence 8667 BP; 1860 A; 3066 C; 2575 G; 1166 T; 0 U; 0 Other;  
Query Match 98.8%; Score 8459.2; DB 11; Length 8667;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 8512; Conservative 1; Mismatches 38; Indels 11; Gaps 2;  
QY 1 CATTGTCGGGCTCCACACAGCTTGTCGACAGACGTGAGGAGCCACAGAGCCCGCTAACCC 60  
DB 115 CATGTCGGGATCCACACAGCTGTGTCGACAGAGCTGAGGAGCCACAGAGCCCGCTAACCC 174  
QY 61 GCCCCACAGCCTTTCTTACCCAGTCAGATCGCCCGGACGACACGAGAGTGGGCTCTT 120  
DB 175 GCCCCACAGCCTTTCTTACCCAGTCAGATCGCCCGGACGACACGAGAGTGGGCTCTT 234  
QY 121 GAGATGACGACACCACTCCCGGACATGATCTCCCACTGTGCGCGGCTTCATATCCA 180  
DB 235 GAGATGACGACACCACTCCCGGACATGATCTCCCACTGTGCGCGGCTTCATATCCA 294  
QY 181 GCCCAGCGCGGAGGSCCTCCCTGCTGTGATGTTCCAGCCCGGGAATGAACGGTCCCA 240  
DB 295 GCCCAGCGCGGAGGSCCTCCCTGCTGTGATGTTCCAGCCCGGGAATGAACGGTCCCA 354  
QY 241 GAGAGTCCACTGCGGCGACAGATCCCACTCATACCTGCGGAGTGGGAAATCAGAGAT 300  
DB 355 GAGAGTCCACTGCGGCGACAGATCCCACTCATACCTGCGGAGTGGGAAATCAGAGAT 414  
QY 301 GAGATTGATTTAAAGCAAGCGCCCTCGGCTGAGAGCTGCTGACCCCTGTGTGACC 360  
DB 415 GAGATTGATTTAAAGCAAGCGCCCTCGGCTGAGAGCTGCTGACCCCTGTGTGACC 474  
QY 361 GTACACCCCTGTGTGCGGACGCGGCTGCGGAGTGTGAAGACTTCAACAAGAACCTGTAG 420  
DB 475 GTACACCCCTGTGTGCGGACGCGGCTGCGGAGTGTGAAGACTTCAACAAGAACCTGTAG 534  
QY 421 CCTGACGGGCAAGTGAACCGGTGTCTCCCGCAGCGCCCGCAGCACTGACCTTGAGCT 480  
DB 535 CCTGACGGGCAAGTGAACCGGTGTCTCCCGCAGCGCCCGCAGCACTGACCTTGAGCT 594  
QY 481 GAGAGTGTGCGGCGACGCGCTGTCCAGAGAGAGTGTATCCAGAACATGACCCGCTGGA 540  
DB 595 GAGAGTGTGCGGCGACGCGCTGTCCAGAGAGAGTGTATCCAGAACATGACCCGCTGGA 654  
QY 541 CGAGAGATCAACATGTGTAGAGAGAGATCTTAAGCTGAAGAAAGCAAGCAAGCT 600  
DB 655 CGAGAGATCAACATGTGTAGAGAGAGATCTTAAGCTGAAGAAAGCAAGCAAGCT 714  
QY 601 GGAGGAGAGAGTGCACAGCGCGCCGAGGCTGAGAGCCGTGTCAACCGCCGCTATCGA 660

DB 715 GGAGGAGAGAGTGCACAGCGCGCCGAGGCTGAGAAACCCGTGTCAACCGCCGCTATCGA 774  
QY 661 GTCCGAAGCACCCGAGGCTGTGAGATCATCTACGAGAACCGGAAAGGCTGAAGC 720  
DB 775 GTCCGAAGCACCCGAGGCTGTGAGATCATCTACGAGAACCGGAAAGGCTGAAGC 834  
QY 721 TGCACATCCGATTTCTGAAAGGCTGTGGGGCCCGAGGTGAGAGTCCGCTGTACAAACGACC 780  
DB 835 TGCACATCCGATTTCTGAAAGGCTGTGGGGCCCGAGGTGAGAGTCCGCTGTACAAACGACC 894  
QY 781 CTCGACACCCGCGAGTATCATGAGAACATCAAAATTAACCGAGCATGCGAAAGACT 840  
DB 895 CTCGACACCCGCGAGTATCATGAGAACATCAAAATTAACCGAGCATGCGAAAGACT 954  
QY 841 AATCTTGATCTTCAAGAGAGAGAAATCAAGCTGTGAAACAATGGAACGAAGTCTGGCA 900  
DB 955 AATCTTGATCTTCAAGAGAGAGAAATCAAGCTGTGAAACAATGGAACGAAGTCTGGCA 1014  
QY 901 GCGCTATGACAGCTCATGAGAGCTTTGAAAAAAAGTGTGAGCGCATGAAAAACAACC 960  
DB 1015 GCGCTATGACAGCTCATGAGAGCTTTGAAAAAAAGTGTGAGCGCATGAAAAACAACC 1074  
QY 961 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGAGATCTACGAAGACAGTTCCTTGAGAT 1020  
DB 1075 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGAGATCTACGAAGACAGTTCCTTGAGAT 1134  
QY 1021 CCGCAAGCAGCCGAGCTTCAGAGGCGCATGCAAGACAGAGTGGCGACGCGGCGAGTGG 1080  
DB 1135 CCGCAAGCAGCCGAGCTTCAGAGGCGCATGCAAGACAGAGTGGCGAGGCGAGTGG 1194  
QY 1081 GCTGTTCATGTGGGCGCGCCGACGAGCAGAGAGGTGTGAGAGATCATGATGGCTCTTC 1140  
DB 1195 GCTGTTCATGTGGGCGCGCCGACGAGCAGACAGAGGTGTGAGAGATCATGATGGCTCTTC 1254  
QY 1141 AGAGCAGAGAAACTGTGAGAGCAATGCGCAAGTGGCGGTGATCCGCCCATGCTGTGA 1200  
DB 1255 AGAGCAGAGAAACTGTGAGAGCAATGCGCAAGTGGCGGTGATCCGCCCATGCTGTGA 1314  
QY 1201 CGAGCTGACACGACGCGCATCAAGTTTCATCAACATGAAACGGGCTTTATGGCCGACCCCAT 1260  
DB 1315 CGAGCTGACACGACGCGCATCAAGTTTCATCAACATGAAACGGGCTTTATGGCCGACCCCAT 1374  
QY 1261 GAAGGTGTCAAAAGACCGCGCAGGTCATGAACATGTGAGTGAACAGAGAGACCTTT 1320  
DB 1375 GAAGGTGTCAAAAGACCGCGCAGGTCATGAACATGTGAGTGAACAGAGAGACCTTT 1434  
QY 1321 CCGGAGAAAGTTTATGACAGCATCCCAAGAACTTTGGCTGATCGATCTCTGTGAGAG 1380  
DB 1435 CCGGAGAAAGTTTATGACAGCATCCCAAGAACTTTGGCTGATCGATCTCTGTGAGAG 1494  
QY 1381 GAAGACATGTGCTGAGTGTGCTCTTATTTACTGATCTTAAGAAATGAGAACTTATTA 1440  
DB 1495 GAAGACATGTGCTGAGTGTGCTCTTATTTACTGATCTTAAGAAATGAGAACTTATTA 1554  
QY 1441 GAGCTGTGTGAGACGGAAGCTTATCGGCGCGCGCAAGAGCCAGCAGCAACAAGCAGCA 1500  
DB 1555 GAGCTGTGTGAGACGGAAGCTTATCGGCGCGCGCAAGAGCCAGCAGCAACAAGCAGCA 1614  
QY 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
DB 1615 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1674  
QY 1561 AGATGAGAAAGAGAAAGAAAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
DB 1675 AGATGAGAAAGAGAAAGAAAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1734  
QY 1621 CGACAAAGAAAGCTCTCTCAAGAGAAAGCAGCAGCAGCAGGAGAGAGAGAGAGAGAG 1680  
DB 1735 CGACAAAGAAAGCTCTCTCAAGAGAAAGCAGCAGCAGCAGGAGAGAGAGAGAGAGAG 1794  
QY 1681 GAAGAGGCTGTGGCTCTCAAAAGGCTCGCAAACTGTCCAAACAGCCAGAGAGAGAGAG 1740  
DB 1795 GAAGAGGCTGTGGCTCTCAAAAGGCTCGCAAACTGTCCAAACAGCCAGAGAGAGAGAG 1854

QY 1741 CCGCATCAACCCGCTCAATGCTAATGAGCCAAACAGGAGAGGCCATCAACCCCAAGA 1800  
DB 1855 CCGCATCAACCCGCTCAATGCTAATGAGCCAAACAGGAGAGGCCATCAACCCCAAGA 1914  
QY 1801 GAGCCCGAGCTGCGCTCCATGAGCTGAATGAGATTCTCGCTGAGCAAGAAAGAAAT 1860  
DB 1915 GAGCCCGAGCTGCGCTCCATGAGCTGAATGAGATTCTCGCTGAGCAAGAAAGAAAT 1974  
QY 1851 GGAACAGCAAGAAAGGCTCTCTGGAACAAGCGCGCAACTGTGCGGCATGCGCGGAT 1920  
DB 1975 GGAACAGCAAGAAAGGCTCTCTGGAACAAGCGCGCAACTGTGCGGCATGCGCGGAT 2034  
QY 1921 GGTGGGCTTCAAGACTGTGTCCAGTGTAAAGACTTCTTCAACTAACAGAAAGGCA 1980  
DB 2035 GGTGGGCTTCAAGACTGTGTGTGCAAGTGAACCTTCTTCAACTAACAGAAAGGCA 2094  
QY 1981 GAACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGAGAAAGAGAGAGAGCGCGC 2040  
DB 2095 GAACCTCGATGAGATCTTGCAGCAGCACAAGCTGAAGATGAGAAAGAGAGAGAGCGCGC 2154  
QY 2041 GAGGAAGAAAGAAAGGCGCGCGCGCGCGCAAGGAGAGGCTGGAATCCGCGCGTGGT 2100  
DB 2155 GAGGAAGAAAGAAAGGCGCGCGCGCGCGCAAGGAGAGGCTGGAATCCGCGCGTGGT 2214  
QY 2101 GAGGATGAGAGATGAGGCGCTCGGCGCTGAGCGGAAATGAGAGAGATGTGAGAGA 2160  
DB 2215 GAGGATGAGAGAGATGAGGCGCTCGGCGCTGAGCGGAAATGAGAGAGATGTGAGAGA 2274  
QY 2161 GGCTGAAGCTTTACATGCTCTCTGAGGAATGAGGTGCCAGAGGGGAATGAGTGGCCACG 2220  
DB 2275 GGCTGAAGCTTTACATGCTCTCTGAGGAATGAGGTGCCAGAGGGGAATGAGTGGCCACG 2334  
QY 2221 CACTGTCAACAACAGCTCAGACACCGAGAGCATTCCTCTCTCTCACTAGAGCGCGCA 2280  
DB 2335 CACTGTCAACAACAGCTCAGACACCGAGAGCATTCCTCTCTCTCACTAGAGCGCGCA 2394  
QY 2281 GGAACAAGGGCAAGATGGGCGCCAGCGCCCAAGCCCTGAGCGCGCAAGCGCCCA 2340  
DB 2395 GGAACAAGGGCAAGATGGGCGCCAGCGCCCAAGCCCTGAGCGCGCAAGCGCCCA 2454  
QY 2341 AGGCCCAACCAACCCCAACAGGAGACATCCGGGCGCCCAATTGAGCCCAACCCGCGCTC 2400  
DB 2455 AGGCCCAACCAACCCCAACAGGAGACATCCGGGCGCCCAATTGAGCCCAACCCGCGCTC 2514  
QY 2401 TGAAGCAACCGAGCGCCCTAAGCCCTCAACAGACCCCTGAGCTCTGCACTCTCTCC 2460  
DB 2515 TGAAGCAACCGAGCGCCCTAAGCCCTCAACAGACCCCTGAGCTCTGCACTCTCTCC 2574  
QY 2461 TGTGTCCTCCAAAGAGAGAGAGAGAGAGAGACCGCAGCAGCGCCCCAGTGAAGAGG 2520  
DB 2575 TGTGTCCTCCAAAGAGAGAGAGAGAGAGAGACCGCAGCAGCGCCCCAGTGAAGAGG 2634  
QY 2521 GAGAGAGCAAGAGCCCGCGCGCTGAGAGACTGGGACATGAGACACAGGAGAGCGGAGGA 2580  
DB 2635 GAGAGAGCAAGAGCCCGCGCGCTGAGAGACTGGGACATGAGACACAGGAGAGCGGAGGA 2694  
QY 2581 GCCCGTCAAGAGAGTGCACGAGAGAGACCGAGGAGGGGCGCGCAAGGGCAAGAGCGC 2640  
DB 2695 GCCCGTCAAGAGAGTGCACGAGAGAGACCGAGGAGGGGCGCGCAAGGGCAAGAGCGC 2754  
QY 2641 GAGAGCGCTGAGGCAACGCGCGAGGAGGCGCTCAAGGCAAGAGAGAGAGGAGGCGAG 2700  
DB 2755 GAGAGCGCTGAGGCAACGCGCGAGGAGGCGCTCAAGGCAAGAGAGAGAGGAGGCGAG 2814  
QY 2701 GCGCAAGGCAACCACTGCGCAAGAGCTCGGCGCGCCCGCAAGCAAGCACTTCAAGTGTAC 2760  
DB 2815 GCGCAAGGCAACCAAGCAAGCAAGCTCGGCGCGCCCGCAAGCAAGCACTTCAAGTGTAC 2874  
QY 2761 CTGAGGTGCAAGAGTGTGAGTGAAGCGGCGGCGCGCAAGAAACCGGCTGTGCCCC 2820  
DB 2875 CTGAGGTGCAAGAGTGTGAGTGAAGCGGCGGCGCGCAAGAAACCGGCTGTGCCCC 2934

QY 2821 AAGGCCAGCTTCTCAACCCGAGCTGCGGACCCCGGCGCAATGCTTCAACCCAGAGGCC 2880  
DB 2935 AAGGCCAGCTTCTCAACCCGAGCTGCGGACCCCGGCGCAATGCTTCAACCCAGAGGCC 2994  
QY 2881 ACTGAACTGAAAGCAAGCTGAAAGCAGCGAGCGCTGCAATCCCCCATCAAGGTGACCA 2940  
DB 2995 ACTGAACTGAAAGCAAGCTGAAAGCAGCGAGCGCTGCAATCCCCCATCAAGGTGACCA 3054  
QY 2941 AGTCATGAGCCCCCGCGGAGAGACGCACTTCCACCAAGCAAGCTTCCCGCAGCCCCAC 3000  
DB 3055 AGTCATGAGCCCCCGCGGAGAGACGCACTTCCACCAAGCAAGCTTCCCGCAGCCCCAC 3114  
QY 3001 GCCACCGCAAAACCTGCAACCGGAGAGAGCGCCCTTCAAGAGCTTGGAGAGCGCCCCG 3060  
DB 3115 GCCACCGCAAAACCTGCAACCGGAGAGAGCGCCCTTCAAGAGCTTGGAGAGCGCCCCG 3174  
QY 3061 GGGCAAGAGCAAGAGCCCGGCAACCCCGCGCAAGAGAGGCTTGGAGCGGAGGCCCA 3120  
DB 3175 GGGCAAGAGCAAGAGCCCGGCAACCCCGCGCAAGAGAGGCTTGGAGCGGAGGCCCA 3234  
QY 3121 GAAGCTGCTGGGAGACCCCTTGTGGAATTCGGGCTGCGCTTCCCGTGCACCCCGCG 3180  
DB 3235 GAAGCTGCTGGGAGACCCCTTGTGGAATTCGGGCTGCGCTTCCCGTGCACCCCGCG 3294  
QY 3181 TGAAGTGAATCAAGGCTTCCCGCATGCCCCGAGACCTTCAAGCTTCTCTTACGCTCAC 3240  
DB 3295 TGAAGTGAATCAAGGCTTCCCGCATGCCCCGAGACCTTCAAGCTTCTCTTACGCTCAC 3354  
QY 3241 TGTGTACCACTGCCCCGAGGCTTCATGACATGCGCGGCGCGCTGCGCGGCGCCAC 3300  
DB 3355 TGTGTACCACTGCCCCGAGGCTTCATGACATGCGCGGCGCGCTGCGCGGCGCCAC 3414  
QY 3301 CACCATCTCAACCCGCTTCCCTCATCTCTCTGCAAGCAACCCAGCGCTTCTGAGAG 3360  
DB 3415 CACCATCTCAACCCGCTTCCCTCATCTCTCTGCAAGCAACCCAGCGCTTCTGAGAG 3474  
QY 3361 GCAAAATAGGTGCATCTTCCCAAGAAATGTGGTTCAGCTTCCAGCGTCCGATCTCAGAG 3420  
DB 3475 GCAAAATAGGTGCATCTTCCCAAGAAATGTGGTTCAGCTTCCAGCGTCCGATCTCAGAG 3534  
QY 3421 TGCCAAAGGCCCCGGGAGGCGCTGTACCATGAGGCTGCGCGCTGATGAGCCCAAAA 3480  
DB 3535 TGCCAAAGGCCCCGGGAGGCGCTGTACCATGAGGCTGCGCGCTGATGAGCCCAAAA 3594  
QY 3481 GCTGCAACCTTTCAGCGAGTGAAGCAGAGCAGCTGTCCCAACGAGGCGAGGCTTGGGC 3540  
DB 3595 GCTGCAACCTTTCAGCGAGTGAAGCAGAGCAGCTGTCCCAACGAGGCGAGGCTTGGGC 3654  
QY 3541 ACCGAGAGGCTTGGGAGTGCACAGCCCAAGAGGCGTCCGTGTGAGAGGAGCAGCTT 3600  
DB 3655 ACCGAGAGGCTTGGGAGTGCACAGCCCAAGAGGCGTCCGTGTGAGAGGAGCAGCTT 3714  
QY 3601 GGGCTCAATTCGGGCGGAGAGATCAACCAAGGCAATTCACAGCAACCGGAGGCGCTCGGA 3660  
DB 3715 GGGCTCAATTCGGGCGGAGAGATCAACCAAGGCAATTCACAGCAACCGGAGGCGCTCGGA 3774  
QY 3661 CAGCGCATCAATACCGCGCTCATCAACCAAGGCAAGCTGAGCTTCTTGAACA 3720  
DB 3775 CAGCGCATCAATACCGCGCTCATCAACCAAGGCAAGCTGAGCTTCTTGAACA 3834  
QY 3721 GGGCAACATCAACAGAGATCATGCGGAGAGACCGGAGTGTGACCCGCGCGCGGA 3780  
DB 3835 GGGCAACATCAACAGAGATCATGCGGAGAGACCGGAGTGTGACCCGCGCGCGGA 3894  
QY 3781 GGAACGCTTGCCTCAAGGCGCAAGTATTAAGAAAGGAGGCGCAAGCTTGTGCTCA 3840  
DB 3895 GGAACGCTTGCCTCAAGGCGCAAGTATTAAGAAAGGAGGCGCAAGCTTGTGCTCA 3954  
QY 3841 TGAAGGTGCAATGTCTGTGACCAAGTGTCCAAAGAGACGCGCAAGAGCTCAGAGACC 3900  
DB 3955 TGAAGGTGCAATGTCTGTGACCAAGTGTCTCAAGAGAGACGCGCAAGAGCTCAGAGACC 4014  
QY 3901 CCCCATGAGACGGCGCGCCCAAGCGCACTTATGACATGATGAGGCGCGGTGGGAG 3960



Db	4015	CCCCATAGACGGCGCCGCCCAAGCGCACTTATGACATGATGAGGGCCCGCTGGGCGAG	4074
Qy	3961	AGCCATCTCTCAGCCAGCATCTGAAAGTCTCATGGGCGGTGCCATCCCGCGGAGCGACA	4020
Db	4075	AGCCATCTCTCAGCCAGCATCTGAAAGTCTCATGGGCGGTGCCATCCCGCGGAGCGACA	4134
Qy	4021	CAGCCCCACCACTTCAAAAGACAGACCAACATCCGCGGGTCCATACACAAGGAGATCC	4080
Db	4135	CAGCCCCACCACTTCAAAAGACAGACCAACATCCGCGGGTCCATACACAAGGAGATCC	4194
Qy	4081	TCGGTCTCACTGAGAGGACAAGAGAGATTAATCTGGTTCGGGAGGCCAAAGTCTTAAAGCG	4140
Db	4195	TCGGTCTCACTGAGAGGACAAGAGAGATTAATCTGGTTCGGGAGGCCAAAGTCTTAAAGCG	4254
Qy	4141	GGAGGGACAGGCTCCGCCCCCAACGGCCCTCACGGGACTTGACCGAGGCTTCAAGAAGCA	4200
Db	4255	GGAGGGACAGGCTCCGCCCCCAACGGCCCTCACGGGACTTGACCGAGGCTTCAAGAAGCA	4314
Qy	4201	GGCCCTGGGGCCCCCTGAAAGCTTAAGACCGGGCCATGAGGGGCTGGTGGCCACGGTGAAGGA	4260
Db	4315	GGCCCTGGGGCCCCCTGAAAGCTTAAGACCGGGCCATGAGGGGCTGGTGGCCACGGTGAAGGA	4374
Qy	4261	GGCGGGCGCTCTCATTCATGAGATTCGGCGGGAAGAGTGGGGACANAGCCCGAGCTGGC	4320
Db	4375	GGCGGGCGCTCTCATTCATGAGATTCGGCGGGAAGAGTGGGGACANAGCCCGAGCTGGC	4434
Qy	4321	CCTGGACCCCGGGGCGGCTCAAGAGAGGGCTCCATACACGAGGGACACCCGCTCAAGTACGA	4380
Db	4435	CCTGGACCCCGGGGCGGCTCAAGAGAGGGCTCCATACACGAGGGACACCCGCTCAAGTACGA	4494
Qy	4381	CACCGGCGGCTCAACCACTGGCTCCAAAAGACAGACGTACGTCTCCCTCATCGGACGCC	4440
Db	4495	CACCGGCGGCTCAACCACTGGCTCCAAAAGACAGACGTACGTCTCCCTCATCGGACGCC	4554
Qy	4441	CGGCGGAGCGTTCCCAACCCGTGACACCGGCTGGAATGTATGGCCGACCGCCGGGCACTGGA	4500
Db	4555	CGGCGGAGCGTTCCCAACCCGTGACACCGGCTGGAATGTATGGCCGACCGCCGGGCACTGGA	4614
Qy	4501	ACGNGCCCTGACAGAGAGAGGACCTGGAAGAAGCCGGCAGGGGACCCGACAGACTGGGGGGG	4560
Db	4615	ACGNGCCCTGACAGAGAGAGGACCTGGAAGAAGCCGGCAGGGGACCCGACAGACTGGGGGGG	4674
Qy	4561	CTCCATTTGCGGCGGGCGCCCGGCTCATTTGATGCTTGAGCTGGGTAAAGCCGCGGACAGCC	4620
Db	4675	CTCCATTTGCGGCGGGCGCCCGGCTCATTTGATGCTTGAGCTGGGCAAGCCGCGGACAGCC	4734
Qy	4621	CTGACCTATGAGGACCAACGGGGGCAACCTTTTGGCGGCCACTTCCACAGAGTTGCGCCGT	4680
Db	4735	CTGACCTATGAGGACCAACGGGGGCAACCTTTTGGCGGCCACTTCCACAGAGTTGCGCCGT	4794
Qy	4681	GACCATGCGGGAGGCCCANCGCGCGGCTGACAGAGAGGAGCCCTTTTCGTTCAGAGAGGATC	4740
Db	4795	GACCATGCGGGAGGCCCANCGCGCGGCTGACAGAGAGGAGCCCTTTTCGTTCAGAGAGGATC	4854
Qy	4741	CCAGAGCCGAAGAAGCTGACGTGACAGGCTCGTAGATGCGCAAGTCCCGCGACAGACACGT	4800
Db	4855	CCAGAGCCGAAGAAGCTGACGTGACAGGCTCGTAGATGCGCAAGTCCCGCGACAGACACGT	4914
Qy	4801	GCCGAGAGCACCCACACCCCATCTGCGCCTTATGAGACACTGTGTTGGGGGCGTAGTGG	4860
Db	4915	GCCGAGAGCACCCACACCCCATCTGCGCCTTATGAGACACTGTGTTGGGGGCGTAGTGG	4974
Qy	4861	CGTGGACCTTGATGCGGACCCACATCCCGCTTGAGACCCCACTTCAATCCCGCGG	4920
Db	4975	CGTGGACCTTGATGCGGACCCACATCCCGCTTGAGACCCCACTTCAATCCCGCGG	5034
Qy	4921	CATCCCTCTGAGACGAGCGCTGCTACTTACCTGCGCCGGAACACTGGGCCCTCAACCCAC	4980
Db	5035	CATCCCTCTGAGACGAGCGCTGCTACTTACCTGCGCCGGAACACTGGGCCCTCAACCCAC	5094
Qy	4981	CTACCCGCACTGTACCCACCTTACCTTATCTCGGGCTTACCCTGACACGCGGCGCTTGA	5040

Db	5095	CTACCCGCACTGTATACCAACCTCTCACTCATCCGGGGCTTCCCGACAAGGGGGGCTTGGAA	5151
Qy	5041	GAACCGGCGAGCCATCATCTCAATGACTACATCACTCGCAGCAGATGCAACAACAACGCG	5100
Db	5155	GAACCGGCGAGCCATCATCTCAATGACTACATCACTCGCAGCAGATGCAACAACAACGCGCG	5214
Qy	5101	CACCGGCATGGCCCAAGGAGCTGATATGCTGAGGGGGCTCTCGGCCCGCCAGTCTCGCT	5160
Db	5215	CACGCGCATGGCCCAAGGAGCTGATATGCTGAGGGGGCTCTCGGCCCGCCAGTCTCGCT	5274
Qy	5161	GGCACTCAACCTACGCTGCGGGGTCCCGAGGCAATCATGACCTGTGCCAATGCGCACACT	5220
Db	5275	GGCACTCAACCTACGCTGCGGGGTCCCGAGGCAATCATGACCTGTGTCCAAATGCGCACACT	5334
Qy	5221	GCTGTGCTCGTGGCCCCCGACAACGAGGACCCCGACGCAACGCGCAATGGAACGCGCTTGGCTA	5280
Db	5335	GCTGTGCTCGTGGCCCCCGACAACGAGGACCCCGACGCAATGGAACGCGCGCTTGGCTA	5394
Qy	5281	CTTCCCAACCGCGCGCCCAAGCCCTTACAGAGCGCGCAACGACGCTTCCCACTTCTCCCAAG	5340
Db	5395	CTTCCCAACCGCGCGCCCAAGCCCTTACAGAGCGCGCAACGACGCTTCCCACTTCTCCCAAG	5454
Qy	5341	AGGTGCAACACACTTTCGCAAAACCAACCAACGAGTCTCGTCCGAGGGGGAGGAGACG	5400
Db	5455	AGGTGCAACACACTTTCGCAAAACCAACCAACGAGTCTCGTCCGAGGGGGAGGAGACG	5514
Qy	5401	GGATCGAGAGCGGAGCCGGGATCGGAGCGGGAATACTCATCTCTCACTCTCAACAGAC	5460
Db	5515	GGATCGAGAGCGGAGCCGGGATCGGAGCGGGAATACTCATCTCTCACTCTCAACAGAC	5574
Qy	5461	GGTGAGACACGACCCATCTTGGAGAATTGGTACAGACGACAGCAGCGCAGCAGCGCAG	5520
Db	5575	GGTGAGACACGACCCATCTTGGAGAATTGGTACAGACGAGCAGCGCAGCAGCGCAG	5634
Qy	5521	CAGCGGGGGGGTGGGGGGCAGCAGACCGCCCGCCCTCCCACTCCCATGCGCAACAGCA	5580
Db	5635	CAGCGGGGGGGTGGGGGGCAGCAGACCGCCCGCCCTCCCACTCCCATGCGCAACAGCA	5694
Qy	5581	CTCGCCCATCTCCCTCTGGAGCCCGAGATGCGCTTCAACAGAGACCAGATGTCTTCAAA	5640
Db	5695	CTCGCCCATCTCCCTCTGGAGCCCGAGATGCGCTTCTACAGAGACCAGATGTGTTCACAA	5754
Qy	5641	CACAGGATGAAAGGATATCATCACCGCTGTGAGGCCAGCAAGCCACGATCTTGAGGTC	5700
Db	5755	CACAGGATGAAAGGATATCATCACCGCTGTGAGGCCAGCAAGCCACGATCTTGAGGTC	5814
Qy	5701	CACCTTCAACTCTCTCAACCGGTTGGCCCAAGTGTGCACATTTCCCACTGTGCACACCACTGCC	5760
Db	5815	CACCTTCAACTCTCTCAACCGGTTGGCCCGAGTGTGCACATTTCCCACTGTGCACACCACTGCC	5874
Qy	5761	ACTGGGGGGGCACTCTGATGGGGGTCTACCCCTCATGGAAGCCGCTCTTGTGCGCCAA	5820
Db	5875	ACTGGGGGGGCACTCTGATGGGGGTCTACCCCTCATGGAAGCCGCTCTTGTGCGCCAA	5934
Qy	5821	GGAAGCCCCCGGGTGCAGCCGAGCGGACCGGACCGGACGAGACACACCGGCAATGCTTCT	5880
Db	5935	GGAAGCCCCCGGGTGCAGCCGAGCGGACCGGACCGGACCGGACGAGACACACCGGCAATGCTTCT	5994
Qy	5881	CGCCAAAGCCCCCAGCCCGCTCGGGCTGAGAGCCCGCTCTTCTCCCAAGCAAGGGCTTGGAA	5940
Db	5995	CGCCAAAGCCCCCAGCCCGCTCGGGCTGAGAGCCCGCTCTTCTCCCAAGCAAGGGCTTGGAA	6054
Qy	5941	GCCCCCGGACCTTATGTGCTCTGTCTTGTGGACACGCAACATGGCGCGCAACCCCTTGGAA	6000
Db	6055	GCCCCCGGACCTTATGTGCTCTGTCTTGTGGACACGCAACATGGCGCGCAACCCCTTGGAA	6114
Qy	6001	GAACCTTGCACCTTCAACACGACGACCGGACCGGACCGGACCGGACCTTGCAGCTTGCAG	6060
Db	6115	GAACCTTGCACCTTCAACACGACGACCGGACCGGACCGGACCGGACCTTGCAGCTTGCAG	6174
Qy	6061	CCGCGACCGGGAAAAAGACTAAAGTAAACCTTTTTCATCTCAGGAATCTGGAATCTCGGTT	6120
Db	6175	CCGCGACCGGGAAAAAGACTAAAGTAAACCTTTTTCATCTCAGGAATCTGGAATCTCGGTT	6234



OY	6121	CTGTGGTTTACA	CGGCAAGACACTTACAGCCCGGAAAGGGGTGGAGGCCGCTCAGCCCTTGAG	6180
Db	6235	TCTGGTTTACA	CGGCAAGACACTTACAGCCCGGAAAGGGGTGGAGGCCGCTTGAG	6294
OY	6181	CTCAACCACTGTGA	CCCAAGAGAGGGGCTCCCAAGACCTGGAGAAGCTCGACAAG	6240
Db	6295	CTCAACCACTGTGA	CCCAAGAGAGGGGCTCCCAAGACCTGGAGAAGCTCGACAAG	6354
OY	6241	CAACTGGAGGGGGA	GCTGCGGCCCAAGCAGCCCGCTGAAGCTTGGCGGGAGGC	6300
Db	6355	CCACTGGAGGGGGA	GCTGCGGCCCAAGCAGCCCGCTGAAGCTTGGCGGGAGGC	6414
OY	6301	CGGCCACCTCCCA	CACTGCGGGCGCTGCGAGAGCGACGGCCCTGTCCAGCCCGCTGCT	6360
Db	6415	CGGCCACCTCCCA	CACTGCGGGCGCTGCGAGAGCGACGGCCCTGTCCAGCCCGCTGCT	6474
OY	6361	CCAGACCGCCCCAG	GGGGTCAACAGCGGGTGTCAACCTTGCGCCAGCATCAG	6420
Db	6475	CCAGACCGCCCCAG	GGGGTCAACAGCGGGTGTCAACCTTGCGCCAGCATCAG	6534
OY	6421	TGAGGTCAATCA	CACAGACTACCCCGGACACCCACAGCAGCTCAGCGCACCTTGCC	6480
Db	6535	TGAGGTCAATCA	CACAGACTACCCCGGACACCCACAGCAGCTCAGCGCACCTTGCC	6594
OY	6481	CGCCCCCTCACT	CTCTCCCTGGGGGACAGTGGCCCGTCTGTGACCTCCGCGCCACCC	6540
Db	6595	CGCCCCCTCACT	CTCTCCCTGGGGGACAGTGGCCCGTCTGTGACCTCCGCGCCACCC	6654
OY	6541	CAGTGACCTTAC	CTCCGCCCGGACCATGTGTGCCCGGCGCGTGTCTCCCGCCACAG	6600
Db	6655	CAGTGACCTTAC	CTCCGCCCGGACCATGTGTGCCCGGCGCGTGTCTCCCGCCACAG	6714
OY	6601	CGAAGGGGGGCA	AGGCTTCCAGAGCCCAACAAAGCTCGGTCTTGGGTGTGTGGA	6660
Db	6715	CGAAGGGGGGCA	AGGCTTCCAGAGCCCAACAAAGCTCGGTCTTGGGTGTGTGGA	6774
OY	6661	CGGATTGAACT	GTGTGTCCCAACGGAGGGGATGACGGAGCCAGGGGCACTCCGGAATGTC	6720
Db	6775	CGGATTGAACT	GTGTGTCCCAACGGAGGGGATGACGGAGCCAGGGGCACTCCGGAATGTC	6834
OY	6721	TGTGTACCCGCT	GTGTACCGGGAATGGGGAAACAGACGGAGCCCAAGAGATGGGCTCAA	6780
Db	6835	TGTGTACCCGCT	GTGTACCGGGAATGGGGAAACAGACGGAGCCCAAGAGATGGGCTCAA	6894
OY	6781	GTCTTCAGGGCA	CACACAGCCGCGCCAGCCTTCTTACAGACTGACCGAGACCACTC	6840
Db	6895	GTCTTCAGGGCA	CACACAGCCGCGCCAGCCTTCTTACAGACTGACCGAGACCACTC	6954
OY	6841	CGCCATGGTCA	AGTCCAAAGAGCAAGATCAACAAAGAGGTGAACCCCAACCGGAA	6900
Db	6955	CGCCATGGTCA	AGTCCAAAGAGCAAGATCAACAAAGAGGTGAACCCCAACCGGAA	7014
OY	6901	TGAGCTGTAAT	AATACAGCCAGCCCTGGAGCGAGATCTTCAATATGCCCCCATCAC	6960
Db	7015	TGAGCTGTAAT	AATACAGCCAGCCCTGGAGCGAGATCTTCAATATGCCCCCATCAC	7074
OY	6961	CGGAAACAGGCT	TATATGACTTATAGAAGCCAGCGGTGTAGAGAACATGCAACCAAT	7020
Db	7075	CGGAAACAGGCT	TATATGACTTATAGAAGCCAGCGGTGTAGAGAACATGCAACCAAT	7134
OY	7021	GGGGCTGGAAGC	CAATATTAAGAAAGGCACTCATGGGTAAATATGACCAAGTGGAGAGTGC	7080
Db	7135	GGGGCTGGAAGC	CAATATTAAGAAAGGCACTCATGGGTAAATATGACCAAGTGGAGAGTGC	7194
OY	7081	CCGCGCGCTCA	GCGGCAATGTCTTTTAAACCTCTGATATGCCAGTCCAGCCCTGCTGC	7140
Db	7195	CCGCGCGCTCA	GCGGCAATGTCTTTTAAACCTCTGATATGCCAGTCCAGCCCTGCTGC	7254
OY	7141	TATGCGCCATA	ACCGTGTGACGAGCGAGATGACACACTTCACTTGCCTGCGAGTGGCGG	7200
Db	7255	TATGCGCCATA	ACCGTGTGACGAGCGAGATGACACACTTCACTTGCCTGCGAGTGGCGG	7314

QY	7201	CGGGAAGGCAAGGCTCTGCGAGACCAGCAGCCGAAAGCCAAAGCTCCCGGCCCGG	7260
Db	7315	CGGGAAGGCAAGGCTCTGCGAGACCAGCAGCCGAAAGCCAAAGCTCCCGGCCCGG	7374
QY	7261	CGTGCATCTGGGGACCGGCAACCCCTGTCTTCTCAGTGCACCTGAGGAGACTGCAA	7320
Db	7375	CGTGCATCTGGGGACCGGCAACCCCTGTCTTCTCAGTGCACCTGAGGAGACTGCAA	7434
QY	7321	CGCGCGGACCGCGTCAACCAACGGGTGGGAGGACAGGCCCTCGTCCGAGTTCCAC	7380
Db	7435	CGCGCGGACCGCGTCAACCAACGGGTGGGAGGACAGGCCCTCGTCCGAGTTCCAC	7494
QY	7381	GCCATTCCCTTACAAACCCCTGATCATGCGGCTGCAGCGGGTGTATGGCTTCCCAAC	7440
Db	7495	GCCATTCCCTTACAAACCCCTGATCATGCGGCTGCAGCGGGTGTATGGCTTCCCAAC	7554
QY	7441	CCCAACCGGGCTTCCCGCGGGCAGCGGGGCCCTCTGCTGGCCCCCACCACCGCTGGACGA	7500
Db	7555	CCCAACCGGGCTTCCCGCGGGCAGCGGGGCCCTCTGCTGGCCCCCACCACCGCTGGACGA	7614
QY	7501	GGAGGCCAAGCAGCTGCTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAGTGA	7560
Db	7615	GGAGGCCAAGCAGCTGCTCTGCTCGCAGTACGAGACACTCTCCGACAGCGAGTGA	7674
QY	7561	AAACAGGGCGGGGGCGGGCGGGCGTGFUEBRS PERLUGTCAAGTCCAGCGACCAAGAA	7620
Db	7675	AAACAGGGCGGGGGCGGGCGGGCGG-----TGTCAAGTCCAGCGACCAAGAA	7724
QY	7621	CGGGCTTCGAGGAGGGGGGGGGCTGCGCATCCGCCCAACCAAGAAAGAGCCCTGAGTC	7680
Db	7725	CGGGCTTCGAGGAGGGGGGGGGCTGCGCATCCGCCCAACCAAGAAAGAGCCCTGAGTC	7784
QY	7681	CGCTGCGCTTCATTCATCTGTCCGTCCAGAGCCGCGACTCTTGTCTTAAAGCTT	7740
Db	7785	CGCTGCGCTTCATTCATCTGTCCGTCCAGAGCCGCGACTCTTGTCTTAAAGCTT	7844
QY	7741	AACTAAGACTCCCGGCCCGGGCTGGCCCTGTGCAGACTTTACTCAGGGATGTTTAACTTG	7800
Db	7845	AACTAAGACTCCCGGCCCGGGCTGGCCCTGTGCAGACTTTACTCAGGGATGTTTAACTTG	7904
QY	7801	GTGCTCGGGAAGGAGGGGAAAGGGCGCGGAGAGGGGCGACCGGACGCTGTGCAGGCAC	7860
Db	7905	GTGCTCGGGAAGGAGGGGAAAGGGCGCGGAGAGGGGCGACCGGACGCTGTGCAGGCAC	7964
QY	7861	ACACAGCGCGGCCAGGGCGGCGCAGGAAACCAAAAGCAGATGACCAGCCTCTCAACGCGCAC	7920
Db	7965	ACACAGCGCGGCCAGGGCGGCGCAGGAAACCAAAAGCAGATGACCAGCCTCTCTCAACGCGCAC	8024
QY	7921	TGCCTCCCCCGAAATGCATTGTGAAACCAAGTCTAAACTGAGCTGCGAGCCCCGGCGCCT	7980
Db	8025	TGCCTCCCCCGAAATGCATTGTGAAACCAAGTCTAAACTGAGCTGCGAGCCCCGGCGCCT	8084
QY	7981	CCCTTCGCGCTCCCATCCGCTTTAGCGCTCTGGAACAGATGAGCGCAGGCGCTGTCAAGCCC	8040
Db	8085	CCCTTCGCGCTCCCATCCGCTTTAGCGCTCTGGAACAGATGAGCGCAGGCGCTGTCTCAAGCCC	8144
QY	8041	CCAGTGCCTGCTTCCGGTTCGCCACAGACTGCCGCCACGACCAAGATGTGTGAAACCAA	8100
Db	8145	CCAGTGCCTGCTTCCGGTTCGCCACAGACTGCCGCCACGACCAAGATGTGTGAAACCAA	8204
QY	8101	GTCAAGCAGAGTGGGCGGACAAAGAGGCGCAGGTGCGGCTGCGGGGGGAAACGATGTCCGA	8160
Db	8205	GTCAAGCAGAGTGGGCGGACAAAGAGGCGCAGGTGCGGCTGCGGGGGGAAACGATGTCCGA	8264
QY	8161	GGACTGACCTGTTTTTTCACATCTGTTGCCGACGCGGTGGGAGAGAAAGCAGATGTGA	8220
Db	8265	GGACTGACCTGTTTTTTCACATCTGTTGCCGACGCGGTGGGAGAGAAAGCAGATGTGA	8324
QY	8221	AATGATGTGCTGTTTACAGGGTATATTTTGTATATCTTCAATGAATTTATTCAGATGTT	8280
Db	8325	AATGATGTGCTGTTTACAGGGTATATTTTGTATATCTTCAATGAATTTATTCAGATGTT	8384
QY	8281	TTACCCAGGAAGGACTTACCCAGATTTACTGCTGTGCTTGTGATCTCTGCTTACCG	8340



Db 661 GTGGAAGCACCGCACCTGGTGTCAATCTATCGACGAGAAACCGAANAAGCTGGAAC 720  
Qy 721 TGCAATCGGATTTCTGAAAGCGCTGGGGCCCCAGGTGAGTGCCTGTACAAACAGCC 780  
Db 721 TGCAATCGGATTTCTGAAAGCGCTGGGGCCCCAGGTGAGTGCCTGTACAAACAGCC 780  
Qy 781 CTCCGACACCGCGCAGTATCATGAGAACATCAAAATTAACAAGCGATGCGAAGAGCT 840  
Db 781 CTCCGACACCGCGCAGTATCATGAGAACATCAAAATTAACAAGCGATGCGAAGAGCT 840  
Qy 841 AATCTTGATCTTCAAGAGGAGATCAAGCTCGGAAAACAATGAAAGCAAAATTTCTGCA 900  
Db 841 AATCTTGATCTTCAAGAGGAGATCAAGCTCGGAAAACAATGAAAGCAAAATTTCTGCA 900  
Qy 901 GCGCTATGACCAAGCTCATGAGAGCGCTTGGAATAAAAGTGAGCGCATGAAAACAACC 960  
Db 901 GCGCTATGACCAAGCTCATGAGAGCGCTTGGAATAAAAGTGAGCGCATGAAAACAACC 960  
Qy 961 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGGAGTACTAGAAAAGCAAGTTCCCTGAGAT 1020  
Db 961 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGGAGTACTAGAAAAGCAAGTTCCCTGAGAT 1020  
Qy 1021 CCGCAAGCAGCGCGCTGCAAGAGCGCATGCAAGCAGGTTGGGCTCAAGCGGCGAGTGG 1080  
Db 1021 CCGCAAGCAGCGCGCTGCAAGAGCGCATGCAAGCAGGTTGGGCTCAAGCGGCGAGTGG 1080  
Qy 1081 GCTGTCAATGTGCGCGCGCCGAGCAGACAGAGGTGTCAGAGATCATGATGCGCTCTC 1140  
Db 1081 GCTGTCAATGTGCGCGCGCCGAGCAGACAGAGGTGTCAGAGATCATGATGCGCTCTC 1140  
Qy 1141 AGAGCAGAGAACTTGAGAGACAGATGCGCACTGCGCTGATCCGCCATGCTGTA 1200  
Db 1141 AGAGCAGAGAACTTGAGAGACAGATGCGCACTGCGCTGATCCGCCATGCTGTA 1200  
Qy 1201 CGAAGCTGACCGCAGCGCATGAAGTTCATCAACATGAAACGGGCTTATGCGCGAACCT 1260  
Db 1201 CGAAGCTGACCGCAGCGCATGAAGTTCATCAACATGAAACGGGCTTATGCGCGAACCT 1260  
Qy 1261 GAAAGTGTACAAAGACCGCCAGGTCAATGAAACATGTGAGTGAAGAGAAAGAACTTT 1320  
Db 1261 GAAAGTGTACAAAGACCGCCAGGTCAATGAAACATGTGAGTGAAGAGAAAGAACTTT 1320  
Qy 1321 CCGGGAAGAACTTCAATGACAGATCCCAAGAACTTTGGCTGATGCAATCTCTGAGAG 1380  
Db 1321 CCGGGAAGAACTTCAATGACAGATCCCAAGAACTTTGGCTGATGCAATCTCTGAGAG 1380  
Qy 1381 GAAAGCAGTGGCTGAGTGGCTCTCTAATTACTGACTAAGAAAGAAATGANAATTATTA 1440  
Db 1381 GAAAGCAGTGGCTGAGTGGCTCTCTAATTACTGACTAAGAAAGAAATGANAATTATTA 1440  
Qy 1441 GAGCTGTGAGACCGAGCTATCGCGCGCGGCAAGAGCCAGCAGCAACACAGACGA 1500  
Db 1441 GAGCTGTGAGACCGAGCTATCGCGCGCGGCAAGAGCCAGCAGCAACACAGACGA 1500  
Qy 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
Db 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
Qy 1561 AGATGAGAAAGAGAGAAAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
Db 1561 AGATGAGAAAGAGAGAAAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
Qy 1621 CGACAAAGAAAGACTTCTCAAGAGAAAGACAGACGACCTCAGGGGAGAGCAACGACGA 1680  
Db 1621 CGACAAAGAAAGACTTCTCAAGAGAAAGACAGACGACCTCAGGGGAGAGCAACGACGA 1680  
Qy 1681 GAAAGAGGCTGTGGCTCTTCAAAGAGCGCGCAAACTGCAACAGCGAGGAAAGCGCAAGG 1740  
Db 1681 GAAAGAGGCTGTGGCTCTTCAAAGAGCGCGCAAACTGCAACAGCGAGGAAAGCGCAAGG 1740  
Qy 1741 CCGCATCAACCGGCTCAATGGCTTAATGAGGCCCAACAGCGAGGCGCATCAACCCCGACGA 1800  
Db 1741 CCGCATCAACCGGCTCAATGGCTTAATGAGGCCCAACAGCGAGGCGCATCAACCCCGACGA 1800

Db 1741 CCGCATCAACCGGCTCAATGGCTTAATGAGGCCCAACAGCGAGGCGCATCAACCCCGACGA 1800  
Qy 1801 GAGCGCCGAGGTGGCTCTCCATGAGACTGAAATGAGAGTTCTCGCTGAGACAGAAAGAAAT 1860  
Db 1801 GAGCGCCGAGGTGGCTCTCCATGAGACTGAAATGAGAGTTCTCGCTGAGACAGAAAGAAAT 1860  
Qy 1861 GGAAGCAGCCAAAGAAAGTCTTCTTGAAACAGCGCGGCAACTGGTGGGCTATGGCCGAT 1920  
Db 1861 GGAAGCAGCCAAAGAAAGTCTTCTTGAAACAGCGCGGCAACTGGTGGGCTATGGCCGAT 1920  
Qy 1921 GGTGGGCTTCAAAGCTGTGTGCAAGTGAAGAACTTCACTTCAACTCAAGAGAGGCA 1980  
Db 1921 GGTGGGCTTCAAAGCTGTGTGCAAGTGAAGAACTTCACTTCAACTCAAGAGAGGCA 1980  
Qy 1981 GAACTTCATGATGATCTTTCAGCAGCAGCAAGCTGAAATGAGAGAGAGAGAGAGAGCGCG 2040  
Db 1981 GAACTTCATGATGATCTTTCAGCAGCAGCAAGCTGAAATGAGAGAGAGAGAGAGAGCGCG 2040  
Qy 2041 GAGAAAGAAAGAAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGGT 2100  
Db 2041 GAGAAAGAAAGAAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGTGGT 2100  
Qy 2101 GAGAGATGAGAGATGAGAGCGGTGCGGGGTGAGCGGAAATGAGAGAGAGATGAGAGGA 2160  
Db 2101 GAGAGATGAGAGATGAGAGCGGTGCGGGGTGAGCGGAAATGAGAGAGAGATGAGAGGA 2160  
Qy 2161 GGTGAAAGCTTTACATGCTCTGAGAAATGAGGTGCCAGAGGGGAAATCAGTGGCCAGC 2220  
Db 2161 GGTGAAAGCTTTACATGCTCTGAGAAATGAGGTGCCAGAGGGGAAATCAGTGGCCAGC 2220  
Qy 2221 CACTGTCAACACAGCTCAGACCTGAGAGCATCCCTCTCTCTCACTGAGGCGCCAA 2280  
Db 2221 CACTGTCAACACAGCTCAGACCTGAGAGCATCCCTCTCTCTCACTGAGGCGCCAA 2280  
Qy 2281 GGAACAAGGGGAGAAATGGGCCCAAGCCCCAGCCACCTTGGGCGCGGAGCGGCGCACCCC 2340  
Db 2281 GGAACAAGGGGAGAAATGGGCCCAAGCCCCAGCCACCTTGGGCGCGGAGCGGCGCACCCC 2340  
Qy 2341 AGGCCACCCACCCACCAAGAGAGCAATCCGGGCGCCCATTTGAGCCACCCCGGCTTC 2400  
Db 2341 AGGCCACCCACCCACCAAGAGAGCAATCCGGGCGCCCATTTGAGCCACCCCGGCTTC 2400  
Qy 2401 TGAAGCCACCGGAGCGCTTGAAGCCCCACACAGACCTCCCATTTGGCTTGCACCTCTCC 2460  
Db 2401 TGAAGCCACCGGAGCGCTTGAAGCCCCACACAGACCTCCCATTTGGCTTGCACCTCTCC 2460  
Qy 2461 TGTGGTCCCAAG 2520  
Db 2461 TGTGGTCCCAAG 2520  
Qy 2521 GAGAGAGCAGAAAGCCCCCGCGGCTGAGAGCTGACAGTGAACAAGGAAAGGCCGAGGA 2580  
Db 2521 GAGAGAGCAGAAAGCCCCCGCGGCTGAGAGCTGACAGTGAACAAGGAAAGGCCGAGGA 2580  
Qy 2581 GCCCGTCAAGAGCGAGTGCACGAGAGAAAGCGAGAGAGGCGCGCAAGGCGCAAGAGAGCG 2640  
Db 2581 GCCCGTCAAGAGCGAGTGCACGAGAGAAAGCGAGAGAGGCGCGCAAGGCGCAAGAGAGCG 2640  
Qy 2641 GGAAGCGCTGAGAGCAGGCGAGGCGGCGGCGCTCAAGGCAAGAGAGAGAGAGAGAGAG 2700  
Db 2641 GGAAGCGCTGAGAGCAGGCGAGGCGGCGGCGCTCAAGGCAAGAGAGAGAGAGAGAGAG 2700  
Qy 2701 CCGCAGGCGCAACCACTGCAAGAGCTCGGGGCGCCCCCAGGACAGCGACTCCAGTGTAC 2760  
Db 2701 CCGCAGGCGCAACCAAGGCAAGAGCTCGGGGCGCCCCCAGGACAGCGACTCCAGTGTAC 2760  
Qy 2761 CTGCAGTGCAAGCAGAGTGAATGAGGCCGAGGCGGCGACAAAGAACCGGCTGTCTCCCG 2820  
Db 2761 CTGCAGTGCAAGCAGAGTGAATGAGGCCGAGGCGGCGGCGACAAAGAACCGGCTGTCTCCCG 2820  
Qy 2821 AAGGCCAGGCTCTCTCAACCGGACTGAGAGACCCCGGGGCAATGCTCAACCCCAAGAGCC 2880  
Db 2821 AAGGCCAGGCTCTCTCAACCGGACTGAGAGACCCCGGGGCAATGCTCAACCCCAAGAGCC 2880

QY 2881 ACTGACCTGAAGCAGCTGAAGCAGGCGGCTGCATCCGCCATCCAGGTCAACAA 2940  
DB 2881 ACTGAGCTGGAAGCACTGAAGCAGGAGGCGCTGCATCCGCCATCCAGGTCAACAA 2940  
QY 2941 AGTCATGAGCCCCCGGAGGAGCGCAGCTCCACCAAGCCAGTCCGCCAGCCCCACCT 3000  
DB 2941 AGTCATGAGCCCCCGGAGGAGCGCAGCTCCACCAAGCCAGTCCGCCAGCCCCACCT 3000  
QY 3001 GCCACCGCAAAAACCTTCAGCCGAGAGCCAGCGCCCTTCAGCAGCTTCGACGAGCCCCG 3060  
DB 3001 GCCACCGCAAAAACCTTCAGCCGAGAGCCAGCGCCCTTCAGCAGCTTCGACGAGCCCCG 3060  
QY 3061 GGGCAAGAGCAGGAGCGCGGCACCCCCGCGCAAGAGAGGCTTCGACGAGCGGACCA 3120  
DB 3061 GGGCAAGAGCAGGAGCGCGGCACCCCCGCGCAAGAGAGGCTTCGACGAGCGGACCA 3120  
QY 3121 GAAAGCTGCTGGGAGACCCCTTGTGTGAACCTTCGGCTTCGCTTCGCCGTGCCCGG 3180  
DB 3121 GAAAGCTGCTGGGAGACCCCTTGTGTGAACCTTCGGCTTCGCTTCGCCGTGCCCGG 3180  
QY 3181 TGAAGGTATCAAGGCTTCCCGGATGCCCCGAGACCCCTCAGCTTCTCTAGCTTCACC 3240  
DB 3181 TGAAGGTATCAAGGCTTCCCGGATGCCCCGAGACCCCTCAGCTTCTCTAGCTTCACC 3240  
QY 3241 TGGTACCCACCTGCGCTGGGAGCTCCATGACACTGCGGCGCGCTGCTGCGCGCCACC 3300  
DB 3241 TGGTACCCACCTGCGCTGGGAGCTCCATGACACTGCGGCGCGCTGCTGCGCGCCACC 3300  
QY 3301 CACCATCTTCAACCCGCTCTCCCTCACTTCTTCGCAAGCACCAGCGTCTTCGAGAG 3360  
DB 3301 CACCATCTTCAACCCGCTCTCCCTCACTTCTTCGCAAGCACCAGCGTCTTCGAGAG 3360  
QY 3361 GCAAAATAGTGCATCTCCCAAGAAATGTCGGTCCAGCTTCAGTCCCGTCACTCAAGACA 3420  
DB 3361 GCAAAATAGTGCATCTCCCAAGAAATGTCGGTCCAGCTTCAGTCCCGTCACTCAAGACA 3420  
QY 3421 TGCAGAGGCCCCGGTGGGCGCTGTCACTAGGGGCTGCCCTGACCATGGAACCCCAAAA 3480  
DB 3421 TGCAGAGGCCCCGGTGGGCGCTGTCACTAGGGGCTGCCCTGACCATGGAACCCCAAAA 3480  
QY 3481 GCTGGACCTTCAAGCGAGTAGAAGCAGAGCAGTGTCCCAAGCGGCGCAAGCTGGGCG 3540  
DB 3481 GCTGGACCTTCAAGCGAGTAGAAGCAGAGCAGTGTCCCAAGCGGCGCAAGCTGGGCG 3540  
QY 3541 ACCGAGAGCTTGGGGGTGCCCAAGCCAGAGGGGTCCGTGTGAAGGAGCACTCT 3600  
DB 3541 ACCGAGAGCTTGGGGGTGCCCAAGCCAGAGGGGTCCGTGTGAAGGAGCACTCT 3600  
QY 3601 GGGCTCAGTTCGGGGGAGAGCATCACAAAGGCAATCCAGACACGGGTCGCTCGGA 3660  
DB 3601 GGGCTCAGTTCGGGGGAGAGCATCACAAAGGCAATCCAGACACGGGTCGCTCGGA 3660  
QY 3661 CAGCGGCATCACTACCGCGGCTTCATCAACCAAGCAGCGCAGCTGAGCTCTGTACAA 3720  
DB 3661 CAGCGGCATCACTACCGCGGCTTCATCAACCAAGCAGCGCAGCTGAGCTCTGTACAA 3720  
QY 3721 GGGCACATCAACAGATCATTCGGCGAGAGCAGCCGAGTCCGTTGACCGCGCGGGA 3780  
DB 3721 GGGCACATCAACAGATCATTCGGCGAGAGCAGCCGAGTCCGTTGACCGCGCGGGA 3780  
QY 3781 GGAAGCGCTGCCAAGGGCGCATCTTACGAAGGCAAGAGGGCAAGCTCTTGTCTTA 3840  
DB 3781 GGAAGCGCTGCCAAGGGCGCATCTTACGAAGGCAAGAGGGCAAGCTCTTGTCTTA 3840  
QY 3841 TGAAGGTGACATGTCTGTGACCCAGTGTCTCAAGAGAGCAGCGCAAGGAGCTCAGAGCC 3900  
DB 3841 TGAAGGTGACATGTCTGTGACCCAGTGTCTCAAGAGAGCAGCGCAAGGAGCTCAGAGCC 3900  
QY 3901 CCCCCATGAGCGGCGCGCCCAAGCGCACTATGACATGATGAGAGGCGCGCTGGGCAAG 3960  
DB 3901 CCCCCATGAGCGGCGCGCCCAAGCGCACTATGACATGATGAGAGGCGCGCTGGGCAAG 3960

QY 3961 AGCCATCTCTCAGCCAGCATCGAAGGTCTATGAGCGGTGCATCCGCGGAGCGGACA 4020  
DB 3961 AGCCATCTCTCAGCCAGCATCGAAGGTCTATGAGCGGTGCATCCGCGGAGCGGACA 4020  
QY 4021 CAGCCCCACACCTTAAAGAGCAGACCAATCCCGGGTCCATCACTCAAGAGATCCC 4080  
DB 4021 CAGCCCCACACCTTAAAGAGCAGACCAATCCCGGGTCCATCACTCAAGAGATCCC 4080  
QY 4081 TCGGTCCTAGTGAAGGACAAGAGGACTACCTGCGTCCGAGAGGCCAAGCTCTTAAAGG 4140  
DB 4081 TCGGTCCTAGTGAAGGACAAGAGGACTACCTGCGTCCGAGAGGCCAAGCTCTTAAAGG 4140  
QY 4141 GAGAGGCAAGCTTCGCGCCCAACCGCTTCACGAGGACTTGAACGAGCTTACAGACGA 4200  
DB 4141 GAGAGGCAAGCTTCGCGCCCAACCGCTTCACGAGGACTTGAACGAGCTTACAGACGA 4200  
QY 4201 GGGCTTGGGCCCCCTGAAGCTGAAGCGGCGCATGAGGGCTGTGTGGCCACGTTAAAGGA 4260  
DB 4201 GGGCTTGGGCCCCCTGAAGCTGAAGCGGCGCATGAGGGCTGTGTGGCCACGTTAAAGGA 4260  
QY 4261 GGGCTTGGGCCCCCTGAAGCTGAAGCGGCGCATGAGGGCTGTGTGGCCACGTTAAAGGA 4320  
DB 4261 GGGCTTGGGCCCCCTGAAGCTGAAGCGGCGCATGAGGGCTGTGTGGCCACGTTAAAGGA 4320  
QY 4321 CCTGGCCCCGCGGCGCTCAAGAGGAGCTCATCAAGCAGAGGACCCCGCTCAAGTACGA 4380  
DB 4321 CCTGGCCCCGCGGCGCTCAAGAGGAGCTCATCAAGCAGAGGACCCCGCTCAAGTACGA 4380  
QY 4381 CACCGGCGGCTCCACCATGAGCTCCAAAAGCAGAGTACGTCCTCTCATTCGGCAGCCC 4440  
DB 4381 CACCGGCGGCTCCACCATGAGCTCCAAAAGCAGAGTACGTCCTCTCATTCGGCAGCCC 4440  
QY 4441 CGGCGGAGAGTTCACACCGGTCGACCGGTCGATGATGATGAGGCGCGCGGCACTGGA 4500  
DB 4441 CGGCGGAGAGTTCACACCGGTCGACCGGTCGATGATGATGAGGCGCGCGGCACTGGA 4500  
QY 4501 ACGTGCCTCTACGAGAGAGCTGGAAGAGCGGCGCAGAGGACCGCAGAGCTTCGGGAGG 4560  
DB 4501 ACGTGCCTCTACGAGAGAGCTGGAAGAGCGGCGCAGAGGACCGCAGAGCTTCGGGAGG 4560  
QY 4561 CTCATTGCGCGCGGCGCGCGCTCATTTGTGCTGAGCTGGGTAAGCGCGGCAAGGCC 4620  
DB 4561 CTCATTGCGCGCGGCGCGCGCTCATTTGTGCTGAGCTGGGTAAGCGCGGCAAGGCC 4620  
QY 4621 CCTGACCTATGAGAGACCAAGGGGCAACCTTTCGCGGCGCACTTCCACGAGGTTCCGCCGT 4680  
DB 4621 CCTGACCTATGAGAGACCAAGGGGCAACCTTTCGCGGCGCACTTCCACGAGGTTCCGCCGT 4680  
QY 4681 GACCATGCGGAGGCCACCGCGCTGCGAGGAGGCAAGCTTTCGTCCAGCAAGGACATC 4740  
DB 4681 GACCATGCGGAGGCCACCGCGCTGCGAGGAGGCAAGCTTTCGTCCAGCAAGGACATC 4740  
QY 4741 CCAAGACCGAAGCTGACGCTGCAAGCTTCGTGAGATGCGCAAGTCCCGCAACGACCT 4800  
DB 4741 CCAAGACCGAAGCTGACGCTGCAAGCTTCGTGAGATGCGCAAGTCCCGCAACGACCT 4800  
QY 4801 GCGCGAGACCAACCAACCAATCTCGGCTATGAGACCTGCTTCGGGGGCTGTGATG 4860  
DB 4801 GCGCGAGACCAACCAACCAATCTCGGCTATGAGACCTGCTTCGGGGGCTGTGATG 4860  
QY 4861 CGTGAAGCTGTATCGAGGCAATCCCTGAGCTTCGACCCCACTCATACCCCGCGG 4920  
DB 4861 CGTGAAGCTGTATCGAGGCAATCCCTGAGCTTCGACCCCACTCATACCCCGCGG 4920  
QY 4921 CATTCCTTGAACGACGCGCTGCTTACTGAGCTTCGAGCTTCGAGCTTCGAGCTTCGAG 4980  
DB 4921 CATTCCTTGAACGACGCGCTGCTTACTGAGCTTCGAGCTTCGAGCTTCGAGCTTCGAG 4980  
QY 4981 CTACCGGACCTGTATCCCACTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 5040  
DB 4981 CTACCGGACCTGTATCCCACTTACCTTACCTTACCTTACCTTACCTTACCTTACCTT 5040  
QY 5041 GAACGGGACAGCATCATCATGACTATCATCTTCGACGAGATGCAACACACGCGC 5100

Db 5041 GAAACGGGAGACATCATCAATGACTATCACTTCGACGAGATGCAACAAAGCGGC 5100  
QY 5101 CACCGCATGCGCCAGCGAGTATATGCTGAGGGGCTCTGCCCCGAGGCTCCGCT 5160  
Db 5101 CACCGCATGCGCCAGCGAGTATATGCTGAGGGGCTCTGCCCCGAGGCTCCGCT 5160  
QY 5161 GGGACTCAACTACGTGCGGGGTCCCCGAGGATCATCGACTGTGCCAATGCCACACT 5220  
Db 5161 GGGACTCAACTACGTGCGGGGTCCCCGAGGATCATCGACTGTGCCAATGCCACACT 5220  
QY 5221 GCGTGTGCTGTGCCCCCGGACACAGGCAACCCAGCCACCCGCAATGAAACCGCTTGCTTA 5280  
Db 5221 GCGTGTGCTGTGCCCCCGGACACAGGCAACCCAGCCACCCGCAATGAAACCGCTTGCTTA 5280  
QY 5281 CTTCGCCACCGGCGCCAGCGCTTCAGACGCGGCAAGCAGCTGCCCACTCTCCCGCAG 5340  
Db 5281 CTTCGCCACCGGCGCCAGCGCTTCAGACGCGGCAAGCAGCTGCCCACTCTCCCGCAG 5340  
QY 5341 AGGTCCAAACAATTGACAAAACAACAACACAGTCTGTCCGAGCGGGAAGAGACG 5400  
Db 5341 AGGTCCAAACAATTGACAAAACAACAACACAGTCTGTCCGAGCGGGAAGAGACG 5400  
QY 5401 GGATGGAAGCGGGACCGGGATCGGGAGCGGGAAGATCCATCTTCACTCCACAGAC 5460  
Db 5401 GGATGGAAGCGGGACCGGGATCGGGAGCGGGAAGATCCATCTTCACTCCACAGAC 5460  
QY 5461 GGTGAGACACGACCCATCTGAGACCTTGATACAGACAGACAGCGGACGAGCGGAG 5520  
Db 5461 GGTGAGACACGACCCATCTGAGACCTTGATACAGACAGACAGCGGAGCGGAG 5520  
QY 5521 CAGCGCGGGGTGAGGAGCAGACGACCGCGCGCTCCCACTCCATGCCACACAGCA 5580  
Db 5521 CAGCGCGGGGTGAGGAGCAGACGACCGCGCGCTCCCACTCCATGCCACACAGCA 5580  
QY 5581 CTGCGCCATCTCCCTCGGACCCAGAGATGCCCTTCAGAGAGAACCCAGTGTCTTCAACA 5640  
Db 5581 CTGCGCCATCTCCCTCGGACCCAGAGATGCCCTTCAGAGAGAACCCAGTGTCTTCAACA 5640  
QY 5641 CACAGGCATGAGGGTATCATCAACGCTGTGAGGCCAGACAGCCACAGTCTTCAAGTGC 5700  
Db 5641 CACAGGCATGAGGGTATCATCAACGCTGTGAGGCCAGACAGCCACAGTCTTCAAGTGC 5700  
QY 5662 CACGAGCATGAGGGTATCATCAACGCTGTGAGGCCAGACAGCCACAGTCTTCAAGTGC 5691  
Db 5662 CACGAGCATGAGGGTATCATCAACGCTGTGAGGCCAGACAGCCACAGTCTTCAAGTGC 5691  
QY 5701 CACCTTCACTCTCAACCGGTTGCGCCAGCTGCCACATTTCCACCTGCCACCACTGCGC 5760  
Db 5701 CACCTTCACTCTCAACCGGTTGCGCCAGCTGCCACATTTCCACCTGCCACCACTGCGC 5760  
QY 5761 ACTGGGCGGCACTCTGATGAGGCTCTAACCCTTACCTCATGAGGCCGCTTGTCTGCCAA 5820  
Db 5761 ACTGGGCGGCACTCTGATGAGGCTCTAACCCTTACCTCATGAGGCCGCTTGTCTGCCAA 5820  
QY 5821 GGAAGCCCCCGGGTCCGCCGCGCAGAGCGGCGCCGAGAGACACCGGCAATGCTTCTCT 5880  
Db 5821 GGAAGCCCCCGGGTCCGCCGCGCAGAGCGGCGCCGAGAGACACCGGCAATGCTTCTCT 5880  
QY 5881 CGCCAAAGCCCCCGGCTCCGGGCTGAGAGCCGCGCTTCTCCCGCAGCAAGGCTCCGA 5940  
Db 5881 CGCCAAAGCCCCCGGCTCCGGGCTGAGAGCCGCGCTTCTCCCGCAGCAAGGCTCCGA 5940  
QY 5941 GCGCCAAAGCCCCCGGCTCCGGGCTGAGAGCCGCGCTTCTCCCGCAGCAAGGCTCCGA 5931  
Db 5941 GCGCCAAAGCCCCCGGCTCCGGGCTGAGAGCCGCGCTTCTCCCGCAGCAAGGCTCCGA 5931  
QY 5991 GCGCCGCGCCCTGATGCTCTGCTCTGAGCAGCAGCAGCATGCGCCGCAACCCCTGCGAA 5991  
Db 5991 GCGCCGCGCCCTGATGCTCTGCTCTGAGCAGCAGCAGCATGCGCCGCAACCCCTGCGAA 5991  
QY 6001 GAACTCTGACCTGACACAGCGACGCGGACCGCGCGGCGCAACGCTGCGGCTCCGA 6060  
Db 6001 GAACTCTGACCTGACACAGCGACGCGGACCGCGCGGCGCAACGCTGCGGCTCCGA 6060  
QY 6061 CCGCGACCTGACCTGACACAGCGACGCGGACCGCGCGGCGCAACGCTGCGGCTCCGA 6051  
Db 6061 CCGCGACCTGACCTGACACAGCGACGCGGACCGCGCGGCGCAACGCTGCGGCTCCGA 6051  
QY 6052 CCGCGACCTGACCTGACACAGCGACGCGGACCGCGCGGCGCAACGCTGCGGCTCCGA 6111  
Db 6052 CCGCGACCTGACCTGACACAGCGACGCGGACCGCGCGGCGCAACGCTGCGGCTCCGA 6111  
QY 6121 TCTGGTTTACCGGAGAGCTTACAGCCCGGAAAGGCTGAGGCCCTGTGTGAG 6180  
Db 6121 TCTGGTTTACCGGAGAGCTTACAGCCCGGAAAGGCTGAGGCCCTGTGTGAG 6180

Db 6112 TCTGGTTTACCGGAGAGCTTACAGCCCGGAAAGGCTGAGGCCCTGTGTGAG 6171  
QY 6181 CTACCCAGTCTGACCCACGACAAAGGGCTCCCAAGCACTGGAAGACTGACAAAG 6240  
Db 6172 CTACCCAGTCTGACCCACGACAAAGGGCTCCCAAGCACTGGAAGACTGACAAAG 6231  
QY 6241 CCACTGGAAGGGAGCTGCGGCCCAAGACAGGCCCCGTGAACTTTGGCGGGAAGC 6300  
Db 6241 CCACTGGAAGGGAGCTGCGGCCCAAGACAGGCCCCGTGAACTTTGGCGGGAAGC 6300  
QY 6232 CCACCTGGAAGGGAGCTGCGGCCCAAGCAAGCCCCGTGAACTTTGGCGGGAAGC 6291  
Db 6232 CCACCTGGAAGGGAGCTGCGGCCCAAGCAAGCCCCGTGAACTTTGGCGGGAAGC 6291  
QY 6301 GCGCCACCTCCCAACCTGCGCGCGCTGAGAGGCGAGGCTGTGTCAGCGCGCTGCT 6360  
Db 6301 GCGCCACCTCCCAACCTGCGCGCGCTGAGAGGCGAGGCTGTGTCAGCGCGCTGCT 6360  
QY 6292 GCGCCACCTCCCAACCTGCGCGCGCTGAGAGGCGAGGCTGTGTCAGCGCGCTGCT 6351  
Db 6292 GCGCCACCTCCCAACCTGCGCGCGCTGAGAGGCGAGGCTGTGTCAGCGCGCTGCT 6351  
QY 6361 CCAGACCGCGCCAGGGGTCAAAAGTCAACAGCGGTGTGACCTTGCGCCAGCAATGAG 6420  
Db 6361 CCAGACCGCGCCAGGGGTCAAAAGTCAACAGCGGTGTGACCTTGCGCCAGCAATGAG 6420  
QY 6352 CCAAGCGCGCCAGGGGTCAAAAGTCAACAGCGGTGTGACCTTGCGCCAGCAATGAG 6411  
Db 6352 CCAAGCGCGCCAGGGGTCAAAAGTCAACAGCGGTGTGACCTTGCGCCAGCAATGAG 6411  
QY 6421 TGAGGTCAATCAACAGACTTACCCGGGACCAACAGAGACTCAGCGCACCTTGCC 6480  
Db 6421 TGAGGTCAATCAACAGACTTACCCGGGACCAACAGAGACTCAGCGCACCTTGCC 6480  
QY 6412 TGAGGTCAATCAACAGACTTACCCGGGACCAACAGAGACTCAGCGCACCTTGCC 6471  
Db 6412 TGAGGTCAATCAACAGACTTACCCGGGACCAACAGAGACTCAGCGCACCTTGCC 6471  
QY 6481 GCGCCCGCTCACTCTTCCGCGGGGCGAGGTGCGCGCTGAGACTCCGCGCGCCAC 6540  
Db 6481 GCGCCCGCTCACTCTTCCGCGGGGCGAGGTGCGCGCTGAGACTCCGCGCGCCAC 6540  
QY 6472 GCGCCCGCTCACTCTTCCGCGGGGCGAGGTGCGCGCTGAGACTCCGCGCGCCAC 6531  
Db 6472 GCGCCCGCTCACTCTTCCGCGGGGCGAGGTGCGCGCTGAGACTCCGCGCGCCAC 6531  
QY 6541 CAGTACCTTACCTCCCGCGCCGAGCAATGTGTCGCCGCGGTGCTCCCCCAGAG 6600  
Db 6541 CAGTACCTTACCTCCCGCGCCGAGCAATGTGTCGCCGCGGTGCTCCCCCAGAG 6600  
QY 6532 CAGTACCTTACCTCCCGCGCCGAGCAATGTGTCGCCGCGGTGCTCCCCCAGAG 6591  
Db 6532 CAGTACCTTACCTCCCGCGCCGAGCAATGTGTCGCCGCGGTGCTCCCCCAGAG 6591  
QY 6601 CGAAGGGGGCAAGAGTCTTCAAGGCCAAACAGAGCTGCTTTGGGTGTGTGAGGA 6660  
Db 6601 CGAAGGGGGCAAGAGTCTTCAAGGCCAAACAGAGCTGCTTTGGGTGTGTGAGGA 6660  
QY 6592 CGAAGGGGGCAAGAGTCTTCAAGGCCAAACAGAGCTGCTTTGGGTGTGTGAGGA 6651  
Db 6592 CGAAGGGGGCAAGAGTCTTCAAGGCCAAACAGAGCTGCTTTGGGTGTGTGAGGA 6651  
QY 6661 CGGTATTGAACTGTGTGTCGCCAGCGAGGGATGACGAGGCGCAGGGGCACTCCGAGTGC 6720  
Db 6661 CGGTATTGAACTGTGTGTCGCCAGCGAGGGATGACGAGGCGCAGGGGCACTCCGAGTGC 6720  
QY 6652 CGGTATTGAACTGTGTGTCGCCAGCGAGGGATGACGAGGCGCAGGGGCACTCCGAGTGC 6711  
Db 6652 CGGTATTGAACTGTGTGTCGCCAGCGAGGGATGACGAGGCGCAGGGGCACTCCGAGTGC 6711  
QY 6721 TGTGTACCCGCTGTCTGTACCGGGATGAGGGAACAGACGAGCCACAGATGAGCTTCA 6780  
Db 6721 TGTGTACCCGCTGTCTGTACCGGGATGAGGGAACAGACGAGCCACAGATGAGCTTCA 6780  
QY 6712 TGTGTACCCGCTGTCTGTACCGGGATGAGGGAACAGACGAGCCACAGATGAGCTTCA 6771  
Db 6712 TGTGTACCCGCTGTCTGTACCGGGATGAGGGAACAGACGAGCCACAGATGAGCTTCA 6771  
QY 6781 GTCTTCAGGCAACACAGCCAGCGCGCTTCTTCAAGAACTGACCGAGAGCACTC 6840  
Db 6781 GTCTTCAGGCAACACAGCCAGCGCGCTTCTTCAAGAACTGACCGAGAGCACTC 6840  
QY 6772 GTCTTCAGGCAACACAGCCAGCGCGCTTCTTCAAGAACTGACCGAGAGCACTC 6831  
Db 6772 GTCTTCAGGCAACACAGCCAGCGCGCTTCTTCAAGAACTGACCGAGAGCACTC 6831  
QY 6841 GCGCAATGTCAAGTCCAAAGAGCAAGAGATCAACAAAGAGCTGAAACCAACCGGAA 6900  
Db 6841 GCGCAATGTCAAGTCCAAAGAGCAAGAGATCAACAAAGAGCTGAAACCAACCGGAA 6900  
QY 6832 GCGCAATGTCAAGTCCAAAGAGCAAGAGATCAACAAAGAGCTGAAACCAACCGGAA 6891  
Db 6832 GCGCAATGTCAAGTCCAAAGAGCAAGAGATCAACAAAGAGCTGAAACCAACCGGAA 6891  
QY 6901 TGAGCTGAATACATATACAGCCAGCTGAGACGAGATCTTCAATATGCCGCGCATGAC 6960  
Db 6901 TGAGCTGAATACATATACAGCCAGCTGAGACGAGATCTTCAATATGCCGCGCATGAC 6960  
QY 6892 TGAGCTGAATACATATACAGCCAGCTGAGACGAGATCTTCAATATGCCGCGCATGAC 6951  
Db 6892 TGAGCTGAATACATATACAGCCAGCTGAGACGAGATCTTCAATATGCCGCGCATGAC 6951  
QY 6961 CGGAAACAGGCTTATGACTTATGAAAGCAGGCGGTGAGGAACTGCAAGCACTAAT 7020  
Db 6961 CGGAAACAGGCTTATGACTTATGAAAGCAGGCGGTGAGGAACTGCAAGCACTAAT 7020  
QY 6952 CGGAAACAGGCTTATGACTTATGAAAGCAGGCGGTGAGGAACTGCAAGCACTAAT 7011  
Db 6952 CGGAAACAGGCTTATGACTTATGAAAGCAGGCGGTGAGGAACTGCAAGCACTAAT 7011  
QY 7021 GGGGCTGGAAGGCAATTAATGAAAGGCACTGATGGGTAAATATGACGAGTGGGAAGATC 7080  
Db 7021 GGGGCTGGAAGGCAATTAATGAAAGGCACTGATGGGTAAATATGACGAGTGGGAAGATC 7080  
QY 7012 GGGGCTGGAAGGCAATTAATGAAAGGCACTGATGGGTAAATATGACGAGTGGGAAGATC 7071  
Db 7012 GGGGCTGGAAGGCAATTAATGAAAGGCACTGATGGGTAAATATGACGAGTGGGAAGATC 7071  
QY 7081 CCGCGCGCTCAGCGCAATGCTTTTAAACCTCTGATGCCAGTGCAGGCTGCCGCTGC 7140  
Db 7081 CCGCGCGCTCAGCGCAATGCTTTTAAACCTCTGATGCCAGTGCAGGCTGCCGCTGC 7140  
QY 7072 CCGCGCGCTCAGCGCAATGCTTTTAAACCTCTGATGCCAGTGCAGGCTGCCGCTGC 7131  
Db 7072 CCGCGCGCTCAGCGCAATGCTTTTAAACCTCTGATGCCAGTGCAGGCTGCCGCTGC 7131  
QY 7141 TATGCCATTAACCGTGTCTGACGAGCGAGTGAACACACTCACCTTGCCAGGTGGCG 7200  
Db 7141 TATGCCATTAACCGTGTCTGACGAGCGAGTGAACACACTCACCTTGCCAGGTGGCG 7200  
QY 7132 TATGCCATTAACCGTGTCTGACGAGCGAGTGAACACACTCACCTTGCCAGGTGGCG 7191  
Db 7132 TATGCCATTAACCGTGTCTGACGAGCGAGTGAACACACTCACCTTGCCAGGTGGCG 7191  
QY 7201 CCGGAAAGGCGCAAGGCTCTGAGCAAGCCAGAGAGCGGAAACCAAGTCCCGGCGCGG 7260  
Db 7201 CCGGAAAGGCGCAAGGCTCTGAGCAAGCCAGAGAGCGGAAACCAAGTCCCGGCGCGG 7260  
QY 7192 CCGGAAAGGCGCAAGGCTCTGAGCAAGCCAGAGAGCGGAAACCAAGTCCCGGCGCGG 7251  
Db 7192 CCGGAAAGGCGCAAGGCTCTGAGCAAGCCAGAGAGCGGAAACCAAGTCCCGGCGCGG 7251

```

OY 7261 CTTGGATCTGGGAGACCGGCACTCTGTCTCTGAGTCACTCGAGAGAGATGCAAA 7320
DB 7252 CTTGGATCTGGGAGACCGGCACTCTGTCTCTGAGTCACTCGAGAGAGATGCAAA 7311
OY 7321 CGGCGGAGACCGGCTCAACCAACCGGCTGTGGAGAGACAGGCGCTGTCCGAGTTCCAC 7380
DB 7312 CGGCGGAGACCGGCTCAACCAACCGGCTGTGGAGAGACAGGCGCTGTCCGAGTTCCAC 7371
OY 7381 GGCATTCCTCTCAACCCCTGATCAATGGGCTGAGGCGGAGTATATGGCTTCCACAC 7440
DB 7372 GGCATTCCTCTCAACCCCTGATCAATGGGCTGAGGCGGAGTATATGGCTTCCACAC 7431
OY 7441 CCCACCGGAGCTCCCGCGGAGACCGGCGCTGTGGGCGGCTCCACACAGCTGGAGACA 7500
DB 7432 CCCACCGGAGCTCCCGCGGAGACCGGCGCTGTGGGCGGCTCCACACAGCTGGAGACA 7491
OY 7501 GGAACCCCAACCACTGCTCTGCTGCAATGAGACACTCTCCGACAGGAGTACTCAG 7560
DB 7492 GGAACCCCAACCACTGCTCTGCTGCAATGAGACACTCTCCGACAGGAGTACTCAG 7551
OY 7561 AACAGGCGGCGGCGGCGGCGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 7620
DB 7552 AACAGGCGGCGGCGGCGGCGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 7599
OY 7621 CGGCGCTGAGAGACCGGCGGCGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 7680
DB 7600 CGGCGCTGAGAGACCGGCGGCGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 7659
OY 7681 CGGCGCTGAGAGACCGGCGGCGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 7740
DB 7660 CGGCGCTGAGAGACCGGCGGCGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 7719
OY 7741 AACTAAGACTCCCGCGGCGGCGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 7800
DB 7720 AACTAAGACTCCCGCGGCGGCGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGG 7779
OY 7801 GTGCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7860
DB 7780 GTGCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7839
OY 7861 ACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7920
DB 7840 ACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7899
OY 7921 TGGCTTCCCGGAGATGATTTGGAAACCAAGTCTTAACTGAGCTCGAGGCGGCGGCTT 7980
DB 7900 TGGCTTCCCGGAGATGATTTGGAAACCAAGTCTTAACTGAGCTCGAGGCGGCGGCTT 7959
OY 7981 CCTCGGCTCCCATCCGCTTATGAGCTCTGAGACAGATGAGACAGGCGGCTGTCCAGGCC 8040
DB 7960 CCTCGGCTCCCATCCGCTTATGAGCTCTGAGACAGATGAGACAGGCGGCTGTCCAGGCC 8019
OY 8041 CCAAGTGGCTGTTCGGATCCCAAGAGTGGCCCAAGAGTCTGAGAACTTCTGGAAACCA 8100
DB 8020 CCAAGTGGCTGTTCGGATCCCAAGAGTGGCCCAAGAGTCTGAGAACTTCTGGAAACCA 8079
OY 8101 GTACAGGCGGAGTGGGCGGAGCAAAAGGCGCAAGTGGCGGCTGTGGGCGGAGAGATGCTCCGA 8160
DB 8080 GTACAGGCGGAGTGGGCGGAGCAAAAGGCGCAAGTGGCGGCTGTGGGCGGAGAGATGCTCCGA 8139
OY 8161 GGAAGTGAATGTTTTTTTTCACACATGTTGCGGACAGCGGAGGAGAAAGGAGATGTA 8220
DB 8140 GGAAGTGAATGTTTTTTTTCACACATGTTGCGGACAGCGGAGGAGAAAGGAGATGTA 8199
OY 8221 AATGATGTGTGGTTTACAGGGTATATTTTGTATATCAATGAATTAATTCAGATGT 8280
DB 8200 AATGATGTGTGGTTTACAGGGTATATTTTGTATATCAATGAATTAATTCAGATGT 8259
OY 8281 TTAACGGAAGAAAGATTAACCAAGTATTAATGCTGTGCTTGTATCTTGTCTTAACCG 8340
DB 8260 TTAACGGAAGAAAGATTAACCAAGTATTAATGCTGTGCTTGTATCTTGTCTTAACCG 8319

```

```

OY 8341 TTCAAGAGCGGTGTGAGAGCGGCAAGTGGGTAGACCCATCACTCGAGAGACCAAGGAGGC 8400
DB 8320 TTCAAGAGCGGTGTGAGAGCGGCAAGTGGGTAGACCCATCACTCGAGAGACCAAGGAGGC 8379
OY 8401 GGGGACTGCTCG-TACGCGCCGCTGTGTCTTCCCTTCTTCTTGGGAGATGA 8459
DB 8380 GGGGACTGCTCG-TACGCGCCGCTGTGTCTTCCCTTCTTCTTGGGAGATGA 8439
OY 8460 ATTGATGCGTATTTCTGTGAGCGGCAATTTGGCGAGGAGGTGTATTTCTCATTTACAC 8519
DB 8440 ATTGATGCGTATTTCTGTGAGCGGCAATTTGGCGAGGAGGTGTATTTCTCATTTACAC 8499
OY 8520 ACGTGTCTTAATTAATAAAGCGAATTAATCTCA 8553
DB 8500 ACGTGTCTTAATTAATAAAGCGAATTAATCTCA 8533

RESULT 6
ACN39603
ID ACN39603 standard; cDNA; 8533 BP.
XX
AC ACN39603;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) cDNA DNA325805, SEQ ID NO:3856.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
XX
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
XX
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
XX
KW central nervous system cancer; bladder cancer; pancreatic cancer;
XX
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
XX
KW chromosome identification; chromosome mapping; gene mapping;
XX
KW gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PX MO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GENTH ) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
XX
PT useful in preparing a medicament for treating or detecting a
XX
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
XX
PT prostate cancer or tumor.
XX
PS Claim 1; SEQ ID NO 3856; 7273bp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
XX
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
XX
CC overexpressed in cancer tissues compared to normal tissues, and may thus
XX
CC serve as effective targets for the diagnosis and treatment of cancer in
XX
CC mammals. The invention also relates to nucleic acid and polypeptide
XX
CC sequences at least 80% identical to the TAT nucleic acids and
XX
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
XX
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
XX
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
XX
CC TAT polypeptide; and methods and compositions for the treatment or
XX
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
XX
CC antibodies, antagonists, binding molecules and compositions are useful
XX
CC for diagnosing or treating a cell proliferative disorder associated with
XX
CC increased TAT expression, particularly cancers such as breast cancer,
XX
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder

```



CC cancer, pancreatic cancer, cervical cancer, cancers of the central  
CC nervous system, melanoma and leukemia. TAT nucleic acids may further be  
CC used as hybridisation probes, in chromosome and gene mapping, in  
CC chromosome identification and in gene therapy. The present sequence  
CC represents a TAT nucleic acid of the invention  
XX

Sequence 8533 BP; 1836 A; 3034 C; 2528 G; 1135 T; 0 U; 0 Other;

Query Match 98.3%; Score 8423.4; DB 13; Length 8533;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 8493; Conservative 3; Mismatches 36; Indels 22; Gaps 3;

QY 1 CATGTCGGGCTCAACAAGCTTGTGGCAACAAGTGAAGGCGCACTGAGCCCGCTAACCC 60  
DB 1 CATGTCGGGATCAACAAGCTTGTGGCAACAAGTGAAGGCGCACTGAGCCCGCTAACCC 60  
QY 61 GCCCCACAGCCCTTCTTACCCGATGAGATCCCGGAGGCAACAAGGATGAGGCTCTCT 120  
DB 61 GCCCCACAGCCCTTCTTACCCGATGAGATCCCGGAGGCAACAAGGATGAGGCTCTCT 120  
QY 121 GAGATGACGACGACACTCCCGGACTATGCTCCCACTGTGCGGGGCTTCATATCCA 180  
DB 121 GAGATGACGACGACACTCCCGGACTATGCTCCCACTGTGCGGGGCTTCATATCCA 180  
QY 181 GCCCCAGCGCGGAGGCGCTCTCTGTGTGTGAATTCAGCCCGGAGATGAACGCTCCA 240  
DB 181 GCCCCAGCGCGGAGGCGCTCTCTGTGTGTGAATTCAGCCCGGAGATGAACGCTCCA 240  
QY 241 GGAGCTCCACCTGGGGCGAGATCCCACTATCCGCGCGGAGCTGGGGAACTCAGAGAT 300  
DB 241 GGAGCTCCACCTGGGGCGAGATCCCACTATCCGCGCGGAGCTGGGGAACTCAGAGAT 300  
QY 301 GGAGTTCAATTGAAACAAAGCGGCTCTGGCTAGAGCTGTGCTGCAACCCCTGCTGCAAC 360  
DB 301 GGAGTTCAATTGAAACAAAGCGGCTCTGGCTAGAGCTGTGCTGCAACCCCTGCTGCAAC 360  
QY 361 GTCAACCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 361 GTCAACCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
QY 421 CCTGACGGGCAAGCTGGAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB 421 CCTGACGGGCAAGCTGGAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
QY 481 GGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
DB 481 GGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
QY 541 CCGAGAGATCAACATGTGTGAGCAGCAGATCTCTAAGCTGAAGAAAGACCAACAGCT 600  
DB 541 CCGAGAGATCAACATGTGTGAGCAGCAGATCTCTAAGCTGAAGAAAGACCAACAGCT 600  
QY 601 GGAGGAGGAGGCTGCAAGCGGCTGCAAGCGGCTGCAAGCGGCTGCAAGCGGCTGCAAGCG 660  
DB 601 GGAGGAGGAGGCTGCAAGCGGCTGCAAGCGGCTGCAAGCGGCTGCAAGCGGCTGCAAGCG 660  
QY 661 GTGCAAGCAGCGCAGCTGTGTGCAAGATCTTACGACGAAACCGGAAAGAGGCTGAAGC 720  
DB 661 GTGCAAGCAGCGCAGCTGTGTGCAAGATCTTACGACGAAACCGGAAAGAGGCTGAAGC 720  
QY 721 TGCAATGATGATTTGTGAAGGCTTGGGGCCCAAGGTGAGGCTTGTGCAACAGCC 780  
DB 721 TGCAATGATGATTTGTGAAGGCTTGGGGCCCAAGGTGAGGCTTGTGCAACAGCC 780  
QY 781 CTCGACACCCGGGAGTATCATGAAGATCAAAATTAACCGAGGATGCGGAAACCT 840  
DB 781 CTCGACACCCGGGAGTATCATGAAGATCAAAATTAACCGAGGATGCGGAAACCT 840  
QY 841 AATCTTGTATTTCAAGAGGAGATCAGGCTGGAACCAATGAGAGGAAATTTCTGCA 900  
DB 841 AATCTTGTATTTCAAGAGGAGATCAGGCTGGAACCAATGAGAGGAAATTTCTGCA 900  
QY 901 GCGCTATGACCACTCATGAGGCTTGGAAAAAAGGTGAGCGCATGAAAAACAACC 960

DB 901 GCGCTATGACCACTCATGAGGCTTGGAGAAAGGTGAGCGCATGAAAAACAACC 960  
QY 961 GCGCGCGCGGCAAG 1020  
DB 961 GCGCGCGCGGCAAG 1020  
QY 1021 CCGCAAGAGCGGAGCTGCAAGAGCGCATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
DB 1021 CCGCAAGAGCGGAGCTGCAAGAGCGCATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080  
QY 1081 GCTGTCCATGTGTGCGCGCGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140  
DB 1081 GCTGTCCATGTGTGCGCGCGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140  
QY 1141 AGAGGAGAGAACTGGAAG 1200  
DB 1141 AGAGGAGAGAACTGGAAG 1200  
QY 1201 GAGCGCTGACGAGCGGATCAAGTTATCATCAATGAACGGGCTTATGCGGACCCGAT 1260  
DB 1201 GAGCGCTGACGAGCGGATCAAGTTATCATCAATGAACGGGCTTATGCGGACCCGAT 1260  
QY 1261 GAAAGTTGACAAAGACCGGCAAGTATGAACATGTGAGTGAAGAGAGAGAGAGAGAGAG 1320  
DB 1261 GAAAGTTGACAAAGACCGGCAAGTATGAACATGTGAGTGAAGAGAGAGAGAGAGAGAG 1320  
QY 1321 CCGGAGAAAGTTCAATGACAGATCCCAAGACTTTGGGCGGATGCGATATCTTGGAGAG 1380  
DB 1321 CCGGAGAAAGTTCAATGACAGATCCCAAGACTTTGGGCGGATGCGATATCTTGGAGAG 1380  
QY 1381 GAAAGCAATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
DB 1381 GAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
QY 1441 GAGCTGTGTGAGCCGAGCTATGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
DB 1441 GAGCTGTGTGAGCCGAGCTATGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500  
QY 1501 GCAGCAG 1560  
DB 1501 GCAGCAG 1560  
QY 1561 AGATGAG 1620  
DB 1561 AGATGAG 1620  
QY 1621 GCAAG 1680  
DB 1621 GCAAG 1680  
QY 1681 GAAAGAGGCTGTGCTCAAGAGCGGCAAACTGCCAAGCGGAGAGAGAGAGAGAGAGAG 1740  
DB 1681 GAAAGAGGCTGTGCTCAAGAGCGGCAAACTGCCAAGCGGAGAGAGAGAGAGAGAGAGAG 1740  
QY 1741 CCGCATCAACCCGCTCAATGAGTGAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800  
DB 1741 CCGCATCAACCCGCTCAATGAGTGAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800  
QY 1801 GAGCGCGAGCTGCTCATGAGAGTGAATGAGATTTCTGCTGAGACAGAGAGAGAGAGAT 1860  
DB 1801 GAGCGCGAGCTGCTCATGAGAGTGAATGAGATTTCTGCTGAGACAGAGAGAGAGAGAT 1860  
QY 1861 GGAAGAGCGGCAAGAAAGTCTCTGGAACAGAGCGGCAAGTGTGCGGATTCGCGCGGAT 1920  
DB 1861 GGAAGAGCGGCAAGAAAGTCTCTGGAACAGAGCGGCAAGTGTGCGGATTCGCGCGGAT 1920  
QY 1921 GATGAGCTCAAGAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980  
DB 1921 GATGAGCTCAAGAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980  
QY 1981 GAACTGATGAGATCTTGCAG 2040





QY	4201	GGCCCTCGGCCCCCTGGAAGCTGGAAGCCGCCCCCTGTAAGGCGCTTGGTGGCCAAAGGTGAAGA	4260
Db	4201	GGCCCTCGGCCCCCTGGAAGCTGGAAGCCGCCCCCTGTAAGGCGCTTGGTGGCCAAAGGTGAAGA	4260
QY	4261	GGCGGGCGCGCTCAATCATAGATCCCGCGAGAGAGCTGCGGCAACAAGCCGAGCTGCC	4320
Db	4261	GGCGGGCGCGCTCAATCATAGATCCCGCGAGAGAGCTGCGGCAACAAGCCGAGCTGCC	4320
QY	4321	CTTGGCCCCGCGCGCTCAAGAGAGGCTCATCAACGAGGGCAACCCGCTCAAGTACGA	4380
Db	4321	CTTGGCCCCGCGCGCTCAAGAGAGGCTCATCAACGAGGGCAACCCGCTCAAGTACGA	4380
QY	4381	CACCGCGCGCTTCACACACTGGCTCCMAAACAACAACGTACGCTTCCCTCATCGGACGCC	4440
Db	4381	CACCGCGCGCTTCACACACTGGCTCCMAAACAACAACGTACGCTTCCCTCATCGGACGCC	4440
QY	4441	CGGCGGAGCTTCCCAACCGGTGCACCCGCTGSAATGTATGGCCGACGCGCGGAGCACTGGA	4500
Db	4441	CGGCGGAGCTTCCCAACCGGTGCACCCGCTGSAATGTATGGCCGACGCGCGGAGCACTGGA	4500
QY	4501	ACGTGCTGTACGAGAGAGCGCTGGAAGCGCGGCAAGGGAACCGCAACAGCTCGGAGG	4560
Db	4501	ACGTGCTGTACGAGAGAGCGCTGGAAGCGCGGCAAGGGAACCGCAACAGCTCGGAGG	4560
QY	4561	CTCCATTGCGCGCGCGCGCGCGCTGATTTGTGCTGAGCTGGGTAAAGCTCGGAGAGGCC	4620
Db	4561	CTCCATTGCGCGCGCGCGCGCGCTGATTTGTGCTGAGCTGGGTAAAGCTCGGAGAGGCC	4620
QY	4621	CCTGACCTATAGAGACCAAGGGGACCCCTTTGGCGGCAACCTCCACAGAGTTGGCCGT	4680
Db	4621	CCTGACCTATAGAGACCAAGGGGACCCCTTTGGCGGCAACCTCCACAGAGTTGGCCGT	4680
QY	4681	GACCATGCGGAGGCCCAAGCGCGCGCTGCAGAGAGGGCAGCTTTGCTCAAGCAAGCATC	4740
Db	4681	GACCATGCGGAGGCCCAAGCGCGCGCTGCAGAGAGGGCAGCTTTGCTCAAGCAAGCATC	4740
QY	4741	CCAGAGCCGAAAGCTGAGTGAAGTCCCAAGTCCCGCACAGCACGT	4800
Db	4741	CCAGAGCCGAAAGCTGAGTGAAGTCCCAAGTCCCGCACAGCACGT	4800
QY	4801	GGCGGAGACCAACCCCATCTGCGCCCTATAGAGCACTGTTGGGGCGTGAAGTG	4860
Db	4801	GGCGGAGACCAACCCCATCTGCGCCCTATAGAGCACTGTTGGGGCGTGAAGTG	4860
QY	4861	CGTGAACCTGTATCGCAGCCACATCCCCCTTGGCCTTGCACCCCATCTCATTCCCGGG	4920
Db	4861	CGTGAACCTGTATCGCAGCCACATCCCCCTTGGCCTTGCACCCCATCTCATTCCCGGG	4920
QY	4921	CATCCCTGTGAGCGAGCGCTGCTACTACCTGCCCCGACACTTGCCCCCAACCCAC	4980
Db	4921	CATCCCTGTGAGCGAGCGCTGCTACTACCTGCCCCGACACTTGCCCCCAACCCAC	4980
QY	4981	CTACCCCGACCTTGAACCCACCTTACCTATCCGCGCTTACCCCAACAGGCGGCGTGA	5040
Db	4981	CTACCCCGACCTTGAACCCACCTTACCTATCCGCGCTTACCCCAACAGGCGGCGTGA	5040
QY	5041	GAAACGGAGACATCATCATGTACTATCACTCGAGAGAGATGCACAACAACGCGC	5100
Db	5041	GAAACGGAGACATCATCATGTACTATCACTCGAGAGAGATGCACAACAACGCGC	5100
QY	5101	CACCGCACATGAGCGAGCGTGTATGTAGAGGGGCTCTCGCCCGGAGTCTCTGCT	5160
Db	5101	CACCGCACATGAGCGAGCGTGTATGTAGAGGGGCTCTCGCCCGGAGTCTCTGCT	5160
QY	5161	GGCACTCAACTACGTGCGGGTCCCCCGAGCATATCACTGTGCCAAGTGCAACCT	5220
Db	5161	GGCACTCAACTACGTGCGGGTCCCCCGAGCATATCACTGTGCCAAGTGCAACCT	5220
QY	5221	GCCTGTGCTGTGCCCCGACACAGGACCCCAAGCACCGCATGTAGCGGCTTGGCTA	5280
Db	5221	GCCTGTGCTGTGCCCCGACACAGGACCCCAAGCACCGCATGTAGCGGCTTGGCTA	5280
QY	5281	CCTCCCCACGCGCCCCGACCTTTACGAGCGCGCACAGCAGCTTCCCACTTCTCCCAAG	5340

Db	5281	CTTCCCAACGCGCCCAAGCCCTTTCAGAGCCGCAACGACGCTCCCACTTCCCAAG	5340
QY	5341	AGCTTCAACAACACTTGACAAACCAACCAACCAACGCTCTCTGTCCGAGCGGAGCGAGACCG	5400
Db	5341	AGCTTCAACAACACTTGACAAACCAACCAACCAACGCTCTCTGTCCGAGCGGAGCGAGACCG	5400
QY	5401	GGATTCGAGAGCGGGACCCGGGATTCGGAGACGGGAAAGTCCATCTTCACTGTCACACACGAC	5460
Db	5401	GGATTCGAGAGCGGGACCCGGGATTCGGAGACGGGAAAGTCCATCTTCACTGTCACACACGAC	5460
QY	5461	GGTGGAGACACGCAACCCATCTGGAGACCTGTGTACAGAGACGAGACGGCGACACGCGGAG	5520
Db	5461	GGTGGAGACACGCAACCCATCTGGAGACCTGTGTACAGAGACGAGACGGCGACACGCGGAG	5520
QY	5521	CAGCGCGCGGGGGTGGGGGACGACGACGCGCCCGCTCCCACTCCCATGCGCCACGACA	5580
Db	5512	CAGCGCGCGGGGGTGGGGGACGACGACGCGCCCGCTCCCACTCCCATGCGCCACGACA	5571
QY	5581	CTGCGCCCATCTCCCTCTGGACCCAGAGATGCCCTTCAGACAGACCCAGTGTCTTCAAA	5640
Db	5572	CTGCGCCCATCTCCCTCTGGACCCAGAGATGCCCTTCAGACAGACCCAGTGTCTTCAAA	5631
QY	5641	CACAGGACATGAGGGGTATCATCAACCGCTGTGGAGCCCAAGCAAGCCCAAGCTTCGAGTTC	5700
Db	5632	CACAGGACATGAGGGGTATCATCAACCGCTGTGGAGCCCAAGCAAGCCCAAGCTTCGAGTTC	5691
QY	5701	CACCTTCAACCTCTCAACCCGTTCCGCCACAGCTGCACATTTCCCACTTCCCAACCACTGCC	5760
Db	5692	CACCTTCAACCTCTCAACCCGTTCCGCCACAGCTGCACATTTCCCACTTCCCAACCACTGCC	5751
QY	5761	ACTGGGCGGACCCCTCGATGGGGGTCTACCCCTATCGAGAGCCGCTCTTGTCTGCCCAA	5820
Db	5752	ACTGGGCGGACCCCTCGATGGGGGTCTACCCCTATCGAGAGCCGCTCTTGTCTGCCCAA	5811
QY	5821	GGAGGCCCCCGGGGTTCGCCGGGCCAGAGCGGGCCCCGAGCAGACACCGGACCATGCTTTCCT	5880
Db	5812	GGAGGCCCCCGGGGTTCGCCGGGCCAGAGCGGGCCCCGAGCAGACACCGGACCATGCTTTCCT	5871
QY	5881	CGCCAAAGCCCCCAGCCCGGCTCCGGGCTGTGAGCCCGGCTCTCCCCCAAGAGGGCTCGGA	5940
Db	5872	CGCCAAAGCCCCCAGCCCGGCTCCGGGCTGTGAGCCCGGCTCTCCCCCAAGAGGGCTCGGA	5931
QY	5941	GCCCCGCGCCCTAGTGCCTCTGTCTGTGGCAACGCAACATTCGACCCCGCAACCCCTGGAA	6000
Db	5932	GCCCCGCGCCCTAGTGCCTCTGTCTGTGGCAACGCAACATTCGACCCCGCAACCCCTGGAA	5991
QY	6001	GAACTCTGACCTTACACACGCGACGCGCGGACCCGCGCGGCGCACCTGCTGTGGCTTGGAA	6060
Db	5992	GAACTCTGACCTTACACACGCGACGCGCGGACCCGCGCGGCGCACCTGCTGTGGCTTGGAA	6051
QY	6061	CCGCGACCGGGGAAAAAGCTCAAAATTAACCTTTTCATCCAGGAACTGGAACTCCGTTTC	6120
Db	6052	CCGCGACCGGGGAAAAAGCTCAAAATTAACCTTTTCATCCAGGAACTGGAACTCCGTTTC	6111
QY	6121	TCTGGGTTTACACGCGGACAGCTTACAGCCCGCGAAGGGGTGGAGCCCGTCAAGCCCTGTGAG	6180
Db	6112	TCTGGGTTTACACGCGGACAGCTTACAGCCCGCGAAGGGGTGGAGCCCGTCAAGCCCTGTGAG	6171
QY	6181	CTCACCCACGTGTGACCAACGACGAAAGGGGCTTCCCAAGACCTTGGAAAGGCTTCGACAAAG	6240
Db	6172	CTCACCCACGTGTGACCAACGACGAAAGGGGCTTCCCAAGACCTTGGAAAGGCTTCGACAAAG	6231
QY	6241	CCACCTGAGGGGGAGGCTGCGCGCCCAAGACGACGAGCCCGTGAAGCTTGGCGGGGAGGC	6300
Db	6232	CCACCTGAGGGGGAGGCTGCGCGCCCAAGACGACGAGCCCGTGAAGCTTGGCGGGGAGGC	6291
QY	6301	CGCCCACTTCCCAACACTGAGCGGCGGCTGTGAGAGCGACGCTCTGTCAAGCCCGCTGCT	6360
Db	6292	CGCCCACTTCCCAACACTGAGCGGCGGCTGTGAGAGCGACGCTCTGTCAAGCCCGCTGCT	6351
QY	6361	CCAAACCGCCCAAGGGGTCAAAGGTCAACAGCGGGTGTCAACCTTGGCCCAACATTCAG	6420

Db	6352	CCAGACCGCCCCAAGGGGTCAAAAGGTCAACAGAGGGGTGTCAACCTGTGGCCACGACATGAC	6411
Qy	6421	TGAGGTCACTCAACAGAGCTACACCCGGCAACAACCCAGAGCTCAAGCGACCCCTGTC	6480
Db	6412	TGAGGTCACTCAACAGAGCTACACCCGGCAACAACCCAGAGCTCAAGCGACCCCTGTC	6471
Qy	6481	CGCCCCCTCTAATCTCTTCCCTGGGGGCCAGCTGCCCCGTCTTGGAATCTCCGCGCCACG	6540
Db	6472	CGCCCCCTCTAATCTCTTCCCTGGGGGCCAGCTGCCCCGTCTTGGAATCTCCGCGCCACG	6531
Qy	6541	CAGTGAACCTTACCTCCGCGCCCCGGGACATAGTATCCCGGCGGTGGCTCCCCACAG	6600
Db	6532	CAGTGAACCTTACCTCCGCGCCCCGGGACATAGTATCCCGGCGGTGGCTCCCCACAG	6591
Qy	6601	CGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGATCGGTCTTGAGTGTGTGAGGA	6660
Db	6592	CGAAGGGGGCAAGAGGTCTCCAGAGCCAAACAAGATCGGTCTTGAGTGTGTGAGGA	6651
Qy	6661	CGGTATTGAACCTGTGTCTCCACCGGAGGGCATGACGAGCCAGGGACCTCCCGAGTGC	6720
Db	6652	CGGTATTGAACCTGTGTCTCCACCGGAGGGCATGACGAGCCAGGGACCTCCCGAGTGC	6711
Qy	6721	TGTGTATCCCGGTGCTGTACCGGAGTGGGGAAACAAGCCGAGGCCAGAGGAATGGGCTCAA	6780
Db	6712	TGTGTATCCCGGTGCTGTACCGGAGTGGGGAAACAAGCCGAGGCCAGAGGAATGGGCTCAA	6771
Qy	6781	GTCTTCAAGGCAACACAGCCAGCCGCGCCAGCTTCTTCAAGCAAGCTGACCGAGACAACTC	6840
Db	6772	GTCTTCAAGGCAACACAGCCAGCCGCGCCAGCTTCTTCAAGCAAGCTGACCGAGACAACTC	6831
Qy	6841	CGCCATGCTCAAGTCCAGAGAGCAAGAGATCAACAAGAGCTGAACAACCCACCGGAA	6900
Db	6832	CGCCATGCTCAAGTCCAGAGAGCAAGAGATCAACAAGAGCTGAACAACCCACCGGAA	6891
Qy	6901	TGAGCCCTGAATATCAATATCAAGCCAGCTCTGGAGCGGAATCTTCAATATGCCCGCATCAC	6960
Db	6892	TGAGCCCTGAATATCAATATCAAGCCAGCTCTGGAGCGGAATCTTCAATATGCCCGCATCAC	6951
Qy	6961	CGGAACAGGCGCTTATGACCTTATGAGAGCGACGGCGGTGCAGGACATGCAACCAACAT	7020
Db	6952	CGGAACAGGCGCTTATGACCTTATGAGAGCGACGGCGGTGCAGGACATGCAACCAACAT	7011
Qy	7021	GGGGCTGAGGCGCAATATTAGAAAAGCATCATGGGTAAATATGACAGTGGGAAAGTCT	7080
Db	7012	GGGGCTGAGGCGCAATATTAGAAAAGCATCATGGGTAAATATGACAGTGGGAAAGTCT	7071
Qy	7081	CCCGCCGCTCAGCGCCCAATGCTTTTAAACCTCTGTGATATCCAGTGCAGGCTGCCGCTGC	7140
Db	7072	CCCGCCGCTCAGCGCCCAATGCTTTTAAACCTCTGTGATATCCAGTGCAGGCTGCCGCTGC	7131
Qy	7141	TATGCGCAATAACCGCTGTGTAGCGGAGCGGAGTGAACAACACTCACTCAGCGAGGTGCGG	7200
Db	7132	TATGCGCAATAACCGCTGTGTAGCGGAGCGGAGTGAACAACACTCACTCAGCGAGGTGCGG	7191
Qy	7201	CGGGAAGGCAAGGATCTCTGGCAAGACCGAGCGGAAAAGCAAGTCCCGGCGCCGGG	7260
Db	7192	CGGGAAGGCAAGGATCTCTGGCAAGACCGAGCGGAAAAGCAAGTCCCGGCGCCGGG	7251
Qy	7261	CTTGGCATCTTGGGAGACGGGCAACCTCTGTCTCTCATGTGCACTTGGAGGGAGCTGCAA	7320
Db	7252	CTTGGCATCTTGGGAGACGGGCAACCTCTGTCTCTCATGTGCACTTGGAGGGAGCTGCAA	7311
Qy	7321	CCGCGCGGACGCGGCTCAACCAACCGCGTGTGGGAAGACAAGGCTCGTCCGAGGTTCCAC	7380
Db	7312	CCGCGCGGACGCGGCTCAACCAACCGCGTGTGGGAAGACAAGGCTCGTCCGAGGTTCCAC	7371
Qy	7381	GCCATTCCCTTACACCCCTGATCATGCGGCTCAGAGCGGATGTCAATGCTTCCCAAC	7440
Db	7372	GCCATTCCCTTACACCCCTGATCATGCGGCTCAGAGCGGATGTCAATGCTTCCCAAC	7431
Qy	7441	CCCAACCGGGCTTCCCGCGGGGACGCGGCGCTCTGTGTGGCCCAACAACGCTTGGAGCA	7500
Db	7432	CCCAACCGGGCTTCCCGCGGGGACGCGGCGCTCTGTGTGGCCCAACAACGCTTGGAGCA	7491

[illegible]

RESULT 7  
ID ACA62249 standard; cDNA; 8686 BP.  
XX ACA62249;  
AC  
XX 12-AUG-2003 (first entry)  
XX  
XX cDNA encoding human nuclear receptor corepressor SMRTE.  
DE  
XX Human; ss; gene; SMRTE; nuclear receptor corepressor; gene therapy;  
KW tissue typing; cancer.  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FT CDS 157..7680  
FT /tag= a  
FT /product= "SMRTE"  
XX  
XX US2003027137-A1.  
XX  
XX 06-FEB-2003.  
XX  
XX 27-MAR-2001; 2001US-00819104.  
XX  
XX 29-MAR-2000; 2000US-0193138P.  
XX  
XX (CHEN/) CHEN J D.  
PA  
XX Chen JD;  
XX  
XX MPI; 2003-466139/44.  
DR P-PSDB; ABU61812.  
XX  
XX New SMRTE proteins and nucleic acids, useful in gene therapy, predictive  
PT medicine, therapeutic or prophylactic treatment, chromosome mapping,  
FT tissue typing and in forensic biology.  
XX  
XX  
XX Claim 2; Page 32-41; 90pp; English.  
XX  
XX The invention relates to an isolated SMRTE nucleic acid molecule. The  
CC nucleic acids are useful in gene therapy, as hybridisation probes for  
CC identifying SMRTE-encoding nucleic acid molecules and as primers for  
CC amplifying of SMRTE nucleic acid molecules. The polypeptides are useful  
CC as immunogens to raise anti-SMRTE antibodies. The SMRTE molecules are  
CC useful as targets for discovering and developing modulating agents to  
CC regulate a variety of cellular processes, in screening assays, in  
CC predictive medicine, in therapeutic or prophylactic treatment, in  
CC chromosome mapping, tissue typing and in forensic identification of a  
CC biological sample. Modulators of SMRTE are useful for treating or  
CC preventing a condition associated with aberrant SMRTE protein or nucleic  
CC acid expression or activity, such as cancer. The present sequence  
CC represents cDNA encoding the human nuclear receptor corepressor SMRTE  
XX  
SQ Sequence 8686 BP; 1881 A; 3069 C; 2562 G; 1174 T; 0 U; 0 Other;  
Query Match 97.2%; Score 8324.6; DB 8; Length 8686;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 8472; Conservative 1; Mismatches 34; Indels 88; Gaps 4;  
QY 1 CATGTCGGGCTCCACACAGCTTGTGGCAGACGCTGGAGGGCCACTGAGCCCGGTACCC 60  
DB 156 CATGTCGGGCTCCACACAGCTTGTGGCAGACGCTGGAGGGCCACTGAGCCCGGTACCC 215  
QY 61 GCCCAGACGCTTTCTTACCCAGTGCAGATCGCCCGAGACGACAGGAGCTGGGCTCT 120  
DB 216 GCCCAGACGCTTTCTTACCCAGTGCAGATCGCCCGAGACGACAGGAGCTGGGCTCT 275  
QY 121 GGAGTACGACGACGACTCCCGGAGCTATGCTCCACGCTGGCGGGCTCATATCCA 180  
DB 276 GGAGTACGACGACGACTCCCGGAGCTATGCTCCACGCTGGCGGGCTCATATCCA 335

QY 181 GCCCAGCGCGGAGAGGCGCTTCCTGCTGTGTAGTTCCAGCCCGGAAATGAACGCTCCA 240  
DB 336 GCCCAGCGCGGAGAGGCGCTTCCTGCTGTGTAGTTCCAGCCCGGAAATGAACGCTCCA 395  
QY 241 GGAGCTCACTCGCGGCGGAGAGTCCCACTCATCTGCGCGGAGCTGGGGAAAGTCAGAGT 300  
DB 396 GGAGCTCACTCGCGGCGGAGAGTCCCACTCATCTGCGCGGAGCTGGGGAAAGTCAGAGT 455  
QY 301 GGAGTTCAATTGAAAGCAAGCGCCCTCGGCTAGAGCTGTGCTGTGACCCCTGCTCGAAC 360  
DB 456 GGAGTTCAATTGAAAGCAAGCGCCCTCGGCTAGAGCTGTGCTGTGACCCCTGCTCGAAC 515  
QY 361 GTACCCCTGCTGTGACCGGCGGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 420  
DB 516 GTACCCCTGCTGTGACCGGCGGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 575  
QY 421 CCGACCGGCGGAGAGTGGAAACCGGTGTCTCCCCGAGCCCGCGACACTGACCCCTGAGT 480  
DB 576 CCGACCGGCGGAGAGTGGAAACCGGTGTCTCCCCGAGCCCGCGACACTGACCCCTGAGT 635  
QY 481 GGAGCTGTGCTGCGGCGGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 540  
DB 636 GGAGCTGTGCTGCGGCGGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 695  
QY 541 CCGAGAGATCAACCATGTGTAGAGACAGACAGATCTTAACTGAAAGAAAGACAGCAACAGT 600  
DB 696 CCGAGAGATCAACCATGTGTAGAGACAGACAGATCTTAACTGAAAGAAAGACAGCAACAGT 755  
QY 601 GGAGAGAGAGGCTGTGCAAGCCCGCGGAGCTGTGAGAGCCCGGTGTCAACCGCGCCATGGA 660  
DB 756 GGAGAGAGAGGCTGTGCAAGCCCGCGGAGCTGTGAGAGCCCGGTGTCAACCGCGCCATGGA 815  
QY 661 GTCGAAGACCGGACGCTGTGTGACAGATCTACGACGAGAACCCGAGAAAGGCTGAAGC 720  
DB 816 GTCGAAGACCGGACGCTGTGTGACAGATCTACGACGAGAACCCGAGAAAGGCTGAAGC 875  
QY 721 TGCACTGTGATTTGTGAAGGCTGTGGGCCCCAGGTGAGCTGCTGTACACACAGCC 780  
DB 876 TGCACTGTGATTTGTGAAGGCTGTGGGCCCCAGGTGAGCTGCTGTACACACAGCC 935  
QY 781 CTCGAGACCCGGGCGGATTCATGAGAACATCAAAATTAACGAGGATGCGGAAAGAGT 840  
DB 936 CTCGAGACCCGGGCGGATTCATGAGAACATCAAAATTAACGAGGATGCGGAAAGAGT 995  
QY 841 AATCTTGTACTTCAAGAGAGGATCAAGCTCGAAGACATGAGAGCAGAACTTGTGCA 900  
DB 996 AATCTTGTACTTCAAGAGAGGATCAAGCTCGAAGACATGAGAGCAGAACTTGTGCA 1055  
QY 901 GCGCTATGACCAAGCTCATGAGAGGCTTGGAAAAAAGGTGAGCGCATCGAAAAACACCC 960  
DB 1056 GCGCTATGACCAAGCTCATGAGAGGCTTGGAAAAAAGGTGAGCGCATCGAAAAACACCC 1115  
QY 961 GCGCGCGGCGGCGAAGAGAGAGAGTGTGCGAGTACTACGAAAGAGGTTCCCTGAGAT 1020  
DB 1116 GCGCGCGGCGGCGAAGAGAGAGAGTGTGCGAGTACTACGAAAGAGGTTCCCTGAGAT 1175  
QY 1021 CCGCAAGACGCGGAGCTGTGACGAGAGCGCATGACAGAGGTTGGGCGACGCGGAGAGTGG 1080  
DB 1176 CCGCAAGACGCGGAGCTGTGACGAGAGCGCATGACAGAGGTTGGGCGACGCGGAGAGTGG 1232  
QY 1081 GCTGTCAATGTGGGCGCGCGGAGCGACGAGAGTGTCAAGATCATGATGGCTCTTC 1140  
DB 1233 GCTGTCAATGTGGGCGCGCGGAGCGAGACGAGAGGTTCAAGATCATGATGGCTCTTC 1292  
QY 1141 AGAGAGAGAGAACTGTGAGAGAGAGATGCGCAGTGTGCGCGGATCCCGCCATGCTGTA 1200  
DB 1293 AGAGAGAGAGAACTGTGAGAGAGAGATGCGCAGTGTGCGCGGATCCCGCCATGCTGTA 1352  
QY 1201 GCAAGCTGACAGACGAGCATCAAGTTTATCAACATGAAACGCGGCTTATGGCCGACCCAT 1260  
DB 1353 GCAAGCTGACAGACGAGCATCAAGTTTATCAACATGAAACGCGGCTTATGGCCGACCCAT 1412

[illegible]

Db	2442	AGGGCAACCCACCCACACACCGAGGACATCCGGCCCCCACTGATGCACCCGGCCTC	2501
Qy	2401	TGAAGCCACCGSAGAGCCCCCTTAGCCCCCAGCAGCACCCTATCGCCCTCTGCACTCTCC	2460
Db	2502	TGAAGCCACCTTAAAGCCCCCTTAGCCCCCAGCAGCACCCTATTCCTTCACTCTCC	2561
Qy	2461	TGTGTCTCCCAAGSAGAGAAAGSAGAGAGACCGCAGCAGCGCCCCCAGTGTAGSAGG	2520
Db	2562	TGTGTCTCCCAAGSAGAGAAAGSAGAGAGACCGCAGCAGCGCCCCCAGTGTAGSAGG	2621
Qy	2521	GGAGSAGCAGAAAGCCCCCGGTGTAGSAGCTGGCAGTGTACACAGSAGAAAGCCGAG	2580
Db	2622	GGAGSAGCAGAAAGCCCCCGGTGTAGSAGCTGGCAGTGTACACAGSAGAAAGCCGAG	2681
Qy	2581	GCCCGTCAAGAGCGAGTGCACGAGSAGAAAGCCGAGSAGGGCCGGCAGAGGCAAGAGCG	2640
Db	2682	GCCCGTCAAGAGCGAGTGCACGAGSAGAAAGCCGAGSAGGGCCGGCAGAGGCAAGAGCG	2741
Qy	2641	GGAGCGCGTGAAGGCCACGCGCCGAGGGGCGCTTCAAGSAGAGAAAGAGGGCGGAG	2700
Db	2742	GGAGCGCGTGAAGGCCACGCGCCGAGGGGCGCTTCAAGSAGAGAAAGAGGGCGGAG	2801
Qy	2701	CGSAGAGGCCACCACTGACCAGAGCTTGGGCGCCCCCAGAGCAGCGCATCTCACTGTAC	2760
Db	2802	CGSAGAGGCCACCACTGAGCCTTGGGCGCCCCCAGAGCAGCGCATCTCACTGTAC	2861
Qy	2761	CTGCAGTGCAGAGAGGTGTAGGTGAGCGCGAGGGCGGAGCAAGAAACGGCGTGTCTCC	2820
Db	2862	CTGCAGTGCAGAGAGGTGTAGGTGAGCGCGAGGGCGGAGCAAGAAACGGCGTGTCTCC	2921
Qy	2821	AAAGCCCAAGCCTCTCAACCCCGACTGTGCGACCCCGGSCCAATGCTCAACCCAGAAAGC	2880
Db	2922	AAAGCCCAAGCCTCTCAACCCCGACTGTGCGACCCCGGSCCAATGCTCAACCCAGAAAGC	2981
Qy	2881	ACTGSACTTGAAGAGCTTGAAGCAGCGAGGGCTGTGCACTCCCCCTTCCAGGTCAACCA	2940
Db	2982	ACTGSACTTGAAGAGCTTGAAGCAGCGAGGGCTGTGCACTCCCCCTTCCAGGTCAACCA	3041
Qy	2941	AGTCAATGAGCCCCCGGGAGGACCGAGCTCCCAACCAAGCAAGCTCCCCAGCCCCCAGC	3000
Db	3042	AGTCAATGAGCCCCCGGGAGGAGCTGAGCTCCCAACCAAGCAAGCTCCCCAGCCCCCAGC	3101
Qy	3001	GCCACCGCAAAACCTGACGCGGAGGAGCGAGCGCCCTCAGCAGCTTGTGAGCAGCCCCG	3060
Db	3102	GCCACCGCAAAACCTGACGCGGAGGAGCGAGCGCCCTCAGCAGCTTGTGAGCAGCCCCG	3161
Qy	3061	GGGCAAGAGCAGAGCCCGGCAACCCCGCCGACAMGAG-----3100	3100
Db	3162	GGGCAAGAGCAGAGCCCGGCAACCCCGCCGACAMGAGAGCTGTGTTCTT3121	3221
Qy	3101	-----GCTTTGGCAGCGGAGGCCAGAGAGCTGCGTGGGAGACCCCTTGTGAGCTTCGG	3156
Db	3222	CCAGGCTTTGGCAGCGGAGGCCAGAGAGCTGCGTGGGAGACCCCTTGTGAGCTTCGG	3281
Qy	3157	CTTGCCTTTCCCGTGGCCCCCGTAGGTATCAAGAGCTTCCCGCATGCCCCCGAGACC	3216
Db	3282	CTTGCCTTTCCCGTGGCCCCCGTAGGTATCAAGAGCTTCCCGCATGCCCCCGAGACC	3341
Qy	3217	CTGAGCTTCTCTACGCTCACCTGTGCAACCATGCGCCCTTGGGCTCTCATGCACTGC	3276
Db	3342	CTGAGCTTCTCTCAAGCTCACTGTGCAACCATGCGCCCTTGGGCTCTCATGCACTGC	3401
Qy	3277	CCGGCCGCTGTGCGCGGCCACCAACCACTTCACCCGCTCTCCCTCATCTCTGTG	3336
Db	3402	CCGGCCGCTGTGCGCGGCCACCAACCACTTCACCCGCTCTCCCTCATCTCTGTG	3461
Qy	3337	CAAGCAACCCAGGCTCTCTGAGAGGCAAAATAGGTGCATCTCCAGAGAAATGTCTGCTCA	3396
Db	3462	CAAGCAACCCAGGCTCTCTGAGAGGCAAAATAGGTGCATCTCCAGAGAAATGTCTGCTCA	3521
Qy	3397	GCTTCACGTCCCGTACTCAGAGCATGCCAAGGCCCGGTGGGCCCTGTGCATGGGGCT	3456

Db 3522 GCTCCAGTCCCTGTACTCAGAGCATGCCAAGGCCCCGGTGGGCTGTGTCACCATG9GGCT 3581  
QY 3457 GCCCTGCCCATGAGACCCCAAAAAGCTGGCACTTTCAGCGGAGTGAAGCAGAGCAAGCT 3516  
Db 3582 GCCCTGCTGCATGAGACCCCAAAAAGCTGGCACTTTCAGCGGAGTGAAGCAGAGCAAGCT 3641  
QY 3517 GTCCCAACGGGGCCAGGCTGGGCCACCCGAGAGGCTTGGGGGTGCCCAACGCCAGAGAGC 3576  
Db 3642 GTCCCAACGGGGCCAGGCTGGGCCACCCGAGAGGCTTGGGGGTGCCCAACGCCAGAGAGC 3701  
QY 3577 GTCCGTGCTGAGAGGAGGAGCAGCTCTGGGCTCATGTTCCGGGGCGGAACATCAACAAAGGCT 3636  
Db 3702 GTCCGTGCTGAGAGGAGGAGCAGCTCTG9GCTCAGTTCCGGGGCGGAACATCAACAAAGGCT 3761  
QY 3637 TCCCAAGCACAGGGGTGCCCTCGAGACGGCCATCAATACCGGGCTCCATCAACCAAGC 3696  
Db 3762 TCCCAAGCACAGGGGTGCCCTCGAGACGGCCATCAATACCGGGCTCCATCAACCAAGC 3821  
QY 3697 CACGCCAGCTGACGTCTGTGTAACAAGGGCAACATCAACAGAGATCATCGGCGAGAGACAGCC 3756  
Db 3822 CACGCCAGCTGACGTCTGTGTAACAAGGGCAACATCAACAGAGATCATCGGCGAGAGACAGCC 3881  
QY 3757 GAGTGGCTTGGACCGGGCGGGGAGAGACAGCTGGCCAAAGGGCCAGTCACTTAAGAAAG 3816  
Db 3882 GAGTGGCTTGGACCGGGCGGGGAGAGACAGCTGGCCAAAGGGCCAGTCACTTAAGAAAG 3941  
QY 3817 CAAGAAAGGCCACAGCTTGTCTCTATAGAGGTGGACATGTCTGTGACCCAGTGTCCAAAG 3876  
Db 3942 CAAGAAAGGCCACAGCTTGTCTCTATAGAGGTGGACATGTCTGTGACCCAGTGTCCAAAG 4001  
QY 3877 GAGCGGCAAGAACAGCTCAGAGACCCGCCATGAGACGGCCGCCCAAGCGCACTTAAG 3936  
Db 4002 GAGCGGCAAGAACAGCTCAGAGACCCGCCATGAGACGGCCGCCCAAGCGCACTTAAG 4061  
QY 3937 CATGATGAGAGGGCGGCTGGGAGAGGCAATCTCTCAAGCAGCATGAGAGGTCTAGGG 3996  
Db 4062 CATGATGAGAGGGCGGCTGGGAGAGGCAATCTCTCAAGCAGCATGAGAGGTCTAGGG 4121  
QY 3997 CCGTGCATATCCCGCCGAGAGCGACAGGCCCCCAACCACTCAAAAGAGCAGCACATCCG 4056  
Db 4122 CCGTGCATATCCCGCCGAGAGCGACAGGCCCCCAACCACTCAAAAGAGCAGCACATCCG 4181  
QY 4057 CCGGTCTCATCACAAGAGGATCTCTGGTCTTACGTGAGGGCACAAGAGGACTTACTTGG 4116  
Db 4182 CCGGTCTCATCACAAGAGGATCTCTGGTCTTACGTGAGGGCACAAGAGGACTTACTTGG 4241  
QY 4117 TCGGGAGGCAAGGCTCTAAAGCGGAGGGGCAAGCTCCGGCCCCACCGGCTCTACAGG 4176  
Db 4242 TCGGGAGGCAAGGCTCTAAAGCGGAGGGGCAAGCTCCGGCCCCACCGGCTCTACAGG 4301  
QY 4177 CCTGACCGAGGCTTCAAGACGCAAGGCTCTGGGCCCTTGAAGCTGAAGCCGGCCATGA 4236  
Db 4302 CCTGACCGAGGCTTCAAGAGCGCAAGGCCCTGGGCCCTTGAAGCTGAAGCCGGCCATGA 4361  
QY 4237 GGGCTGTGGTCAACGTGAGAGAGCGGGCCGCTTCATTCATGAGATCCGCGAGAGA 4296  
Db 4362 GGGCTGTGGTCAACGTGAGAGAGCGGGCCGCTTCATTCATGAGATCCGCGAGAGA 4421  
QY 4297 GCTGCGGACACGCGCGAGCTGCGCTGGCCCGGGCGCGCTCAAGAGGGGCTCCATAC 4356  
Db 4422 GCTGCGGACACGCGCGAGCTGCGCTGGCCCGGGCGCGCTCAAGAGGGGCTCCATAC 4481  
QY 4357 GCAGGCGACCCCGCTCAAGTACGACAGCGGCGGCTTCACACATGAGCTCCAAAAACACA 4416  
Db 4482 GCAGGCGACCCCGCTCAAGTACGACAGCGGCGGCTTCACACATGAGCTCCAAAAACACA 4541  
QY 4417 CGTACGCTTCCATATCGGAGCGCCCGGCGGAGCTTCCACCCGTCACCCGCTGAGTGT 4476  
Db 4542 CGTACGCTTCCATATCGGAGCGCCCGGCGGAGCTTCCACCCGTCACCCGCTGAGTGT 4601  
QY 4477 GATGCGCAGACGCGGGGCACTGGAGAGTGGCTGTAACGAGGAGACCTGAAGACCGGCG 4536  
Db 4602 GATGCGCAGACGCGGGGCACTGGAGAGTGGCTGTAACGAGGAGACCTGAAGACCGGCG 4661

QY 4537 AGGAGCCGCAAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCGGTCAATTGTGCTGA 4596  
Db 4662 AGGAGCCGCAAGCAGCTCGGGGGGCTCCATTGCGCGCGGCGCCCGGTCAATTGTGCTGA 4721  
QY 4557 GCTGGTAAAGCCGCGGAGAGACCCCTTGAACCTATAGAGACCAACGGGGGACCTTTTGGCG 4656  
Db 4722 GCTGGTAAAGCCGCGGAGAGACCCCTTGAACCTATAGAGACCAACGGGGGACCTTTTGGCG 4781  
QY 4657 CCACCTCCACGAGGTTGCGCCGTTGACCATGCGGGAGGCCAGCGCGGCTTGCAGAGAG 4716  
Db 4782 CCACCTCCACGAGGTTGCGCCGTTGACCATGCGGGAGGCCAGCGCGGCTTGCAGAGAG 4841  
QY 4717 CAGCTTTTGTTCACAGAGGCAATCCACAGACCGAAAGGTGAGTGAAGCTCTGTAGAT 4776  
Db 4842 CAGCTTTTGTTCACAGAGGCAATCCACAGACCGAAAGGTGAGTGAAGCTCTGTAGAT 4901  
QY 4777 CGCCAAAGTCCCGGACAGACACCTGTCCGAGACCAACCCCACTGTGGCCCTTAAG 4836  
Db 4902 CGCCAAAGTCCCGGACAGACACCTGTCCGAGACCAACCCCACTGTGGCCCTTAAG 4961  
QY 4837 GCACCTGCTTGGGGCGTGAAGTGGGTGAACCTGTATGACAGCCACATCCCGTGGCTT 4896  
Db 4962 GCACCTGCTTGGGGCGTGAAGTGGGTGAACCTGTATGACAGCCACATCCCGTGGCTT 5021  
QY 4897 GCACCCCACTTCATACCCCGGGGATCCCTGTGAGCGAGCGGCTGCTTACTACTTGGCC 4956  
Db 5022 GCACCCCACTTCATACCCCGGGGATCCCTGTGAGCGAGCGGCTGCTTACTACTTGGCC 5081  
QY 4957 CCGAATCTTGGCCCCCAACCCCACTTACCGGCACTTGTATCCCACTTACCTCAATCCGG 5016  
Db 5082 CCGAATCTTGGCCCCCAACCCCACTTACCGGCACTTGTATCCCACTTACCTCAATCCGG 5141  
QY 5017 CTACCCCGACACGGGGGCGTGGAGAACGGGACAGACATCAATGACTACATCACTTC 5076  
Db 5142 CTACCCCGACACGGGGGCGTGGAGAACGGGACAGACATCAATGACTACATCACTTC 5201  
QY 5077 GCAGAGATGACACCAACAGGCCACCGGCACCGGAGCGGAGCTGATATGAGAGG 5136  
Db 5202 GCAGAGATGACACCAACAGGCCACCGGCACCGGAGCGGAGCTGATATGAGAGG 5261  
QY 5137 CCTTCGGCCCCGAGTCTCTGGTGGCACTCACTACGCTGGGTCCCCGAGGATCAT 5196  
Db 5262 CCTTCGGCCCCGAGTCTCTGGTGGCACTCACTACGCTGGGTCCCCGAGGATCAT 5321  
QY 5197 GCACCTGTCCCAAGGACCACTGCTGTGTGCTGTGCCCCGACACAGGAGCCCAAGC 5256  
Db 5322 GCACCTGTCCCAAGTGCACACTGTGCTGTGTGCTGTGCCCCGACACAGGAGCCCAAGC 5381  
QY 5257 CACCGCATGAGACCGCTTGTGCTTACCTTCCACCGGCGCCAGCCCTTCAAGAGCGGCA 5316  
Db 5382 CACCGCATGAGACCGCTTGTGCTTACCTTCCACCGGCGCCAGCCCTTCAAGAGCGGCA 5441  
QY 5317 CAGAGCTTCCCACTTCTCCCGAGAGGTCCAAACAATTGGAACAAAACACACACGTC 5376  
Db 5442 CAGAGCTTCCCACTTCTCCCGAGAGGTCCAAACAATTGGAACAAAACACACACGTC 5501  
QY 5377 CTCCGTCCGAGCGGAGCAGACCGGGATCGAGAGCGGAGACCGGATTCGGAGCGGGAAA 5436  
Db 5502 CTCCGTCCGAGCGGAGCAGACCGGGATTCGAGAGCGGAGACCGGATTCGGAGCGGGAAA 5561  
QY 5437 GTCCATCTCTCAAGTCCACACGACGAGTGAAGCAGCACCTTGTGAGACCTGTGTAAGA 5496  
Db 5562 GTCCATCTCTCAAGTCCACACGAGGATGAGACGCAACCACTTGTGAGACCTGTGTAAGA 5621  
QY 5497 GCAGAGCAGCGGACAGACCGGCGAGAGCGGGGGTGGGGGACAGACAGCGCGCGCGC 5556  
Db 5622 GCAGAGCAGCGGACAGACCGGCGAGAGCGGGGGTGGGGGACAGACAGCGCGCGCGC 5681  
QY 5557 CTCCCATCTCCATGCGCACAGCACTGCGCCATCTCCCTCGAGCCAGAGATGCCCTCA 5616  
Db 5682 CTCCCATCTCCATGCGCACAGCACTGCGCCATCTCCCTCGAGCCAGAGATGCCCTCA 5741



QY	5617	GCAGAGACCCAGTGTGTCTTCAACAACAAGCATGAAGGTATCATCAACCGCTGTGAAGCC	5676
Db	5742	GCAGAGACCCAGTGTGTCTTCAACAACAAGCATGAAGGTATCATCAACCGCTGTGAAGCC	5801
QY	5677	CAGCAAGGCCACGGTCTGAGGTCCACCTCCACCGTGGCCGAGCTGCAC	5736
Db	5802	CAGCAAGGCCACGGTCTGAGGTCCACCTCCACCGTGGCCGAGCTGCAC	5861
QY	5737	ATTCGCACCTGCACACCCATGCGCCACTGTGGGCGCACCTTCGATGGGGTCTACCTTACCT	5796
Db	5862	ATTCGCACCTGCACACCCATGCGCCACTGTGGGCGCACCTTCGATGGGGTCTACCTTACCT	5921
QY	5797	CATGGAAGCCCGCTCTTGTCTGTCCCAAGAGAGGCCCGCGGTCGCGGCAAGAGCGGCCCG	5856
Db	5922	CATGGAAGCCCGCTCTTGTCTGTCCCAAGAGAGGCCCGCGGTCGCGGCAAGAGCGGCCCG	5981
QY	5857	AGCAGACACCGGGCCATGCTCTTCCTGCGCAAGCCCCCAGCCCGCTCCGGGGCTGAAGCCCG	5916
Db	5982	AGCAGACACCGGGCCATGCTCTTCCTGCGCAAGCCCCCAGCCCGCTCCGGGGCTGAAGCCCG	6041
QY	5917	CTCTCCCCCAGCAAGGGCTCGAGAGCCCGGCCCTTAGTGTGCTCTGTCTCTGACACGC	5976
Db	6042	CTCTCCCCCAGCAAGGGCTCGAGAGCCCGGCCCTTAGTGTGCTCTGTCTCTGACACGC	6101
QY	5977	CACCATGCGCCGCAACCCCTGCGAAGAACCTTCGACCAACGCGCAGCCCGAACCCGCC	6036
Db	6102	CACCATGCGCCGCAACCCCTGCGAAGAACCTTCGACCAACGCGCAGCCCGAACCCGCC	6161
QY	6037	GGCGGCACCTGTCTGTGGCTTGGACACCGGCACCCGGGAAAAGCTCAAGTAAACCTTTTC	6096
Db	6162	GGCGGCACCTGTCTGTGGCTTGGACACCGGCACCCGGGAAAAGCTCAAGTAAACCTTTTC	6221
QY	6097	CATCAGAGAACTGGAACTCCGGTCTCTGGGGTTTACAGCGGAGCAGCTACAGCCCGAAG	6156
Db	6222	CATCAGAGAACTGGAACTCCGGTCTCTGGGGTTTACAGCGGAGCAGCTACAGCCCGAAG	6281
QY	6157	GGTGAGAGCCGTCAGCCCTGTGAAGTCAACCAAGTCTGACCCACGACAAAGGGGCTCCCAA	6216
Db	6282	GGTGAGAGCCGTCAGCCCTGTGAAGTCAACCAAGTCTGACCCACGACAAAGGGGCTCCCAA	6341
QY	6217	GCACCTGGAAAGCTCCGACAAAGACCACTTGAAGGGGAGCTGCGGCCCAAGACCGACG	6276
Db	6342	GCACCTGGAAAGCTCCGACAAAGACCACTTGAAGGGGAGCTGCGGCCCAAGACCGACG	6401
QY	6277	CCCGGTGAAGCTTGGCGGGGAGGGCGCCCACTCCACACTGGGGCGCGCTGAGAG	6336
Db	6402	CCCGGTGAAGCTTGGCGGGGAGGGCGCCCACTCCACACTGGGGCGCGCTGAGAG	6461
QY	6337	CCAGCCCTCGTCCAGCCCGGCTGTCCAGACCGGCCAGGGGTCAAAAGTCAACAGACGGGT	6396
Db	6462	CCAGCCCTCGTCCAGCCCGGCTGTGTCCAGACCGGCCAGGGGTCAAAAGTCAACAGACGGGT	6521
QY	6397	GGTCAACCTTGAGCCAGCAATCAGTGAAGTCATCAACAGAGCTAACACCCGGCACACCC	6456
Db	6522	GGTCAACCTTGAGCCAGCAATCAGTGAAGTCATCAACAGAGCTAACACCCGGCACACCC	6581
QY	6457	ACACAGAGTCAAGGCCACCCCTTGGCCCCCCCCCTTACTCTTCCCTGGGGCACAATGCC	6516
Db	6582	ACACAGAGTCAAGGCCACCCCTTGGCCCCCCCCCTTACTCTTCCCTGGGGCACAATGCC	6641
QY	6517	CGTCTTGGAACCTCGAGCGGCCCAACCAAGTGAACCTTACTCCCGGCCCGGACATGGTGC	6701
Db	6642	CGTCTTGGAACCTCGAGCGGCCCAACCAAGTGAACCTTACTCCCGGCCCGGACATGGTGC	6701
QY	6577	CCCGAGCCCGTGGCTCCCCCAACAGCGAAGGGGGCAAGAGTCTTCAGAGCCAAACAAGAC	6636
Db	6702	CCCGAGCCCGTGGCTCCCCCAACAGCGAAGGGGGCAAGAGTCTTCAGAGCCAAACAAGAC	6761
QY	6637	GTCGGTCTTGGGTGTGTGTGAAGACGGATTGAACCTGTGTCCCCACCGAGGGGCAATGAC	6696
Db	6762	GTCGGTCTTGGGTGTGTGTGAAGACGGATTGAACCTGTGTCCCCACCGAGGGGCAATGAC	6821
QY	6697	GGAGCCAGAGGACCTCCCGAGTGTGTTTACCCCGCTGTATCCGGGATGGGGAAACAGAC	6756

Db	6822	GGAGCCAGGGGACCTCCGGAGTGTGTGTACCTCCCTGTGTACCGGGATGGGGAAAGAC	6881
OY	6757	GGAGCCCGCAGCAGATYGGGCTTCAAGTCTCCAGGGCAACACACGACGCGCAGGCTTCTT	6816
Db	6882	GGAGCCCGCAGCAGATYGGGCTTCAAGTCTCCAGGGCAACACGACGCGCAGGCTTCTT	6941
OY	6817	CAGCAAGCTGACCGAGAGCAACTCCGCAATGTCTCAAGTCCAGAGACAGATCAACAA	6876
Db	6942	CAGCAAGCTGACCGAGAGCAACTCCGCAATGTCTCAAGTCCAGAGACAGATCAACAA	7001
OY	6877	GAAGCTGAACCCCAACACCGGGAATGAGCTGAATACAAATATCAGCCAGCTTGGGACGA	6936
Db	7002	GAAGCTGAACCCCAACACCGGGAATGAGCTGAATACAAATATCAGCCAGCTTGGGACGA	7061
OY	6937	GATCTTCAATATGCGCGCATCACCGGAAAGGCTTATGACTATATGAAGCCAGGCGGT	6996
Db	7062	GATCTTCAATATGCGCGCATCACCGGAAAGGCTTATGACTATATGAAGCCAGGCGGT	7121
OY	6997	GCAGAGAAATGTCCAGACCAACATGTGGGGCTGGAGGCCATATATGAAAGGCACTCATGG	7056
Db	7122	GCAGAGAAATGTCCAGACCAACATGTGGGGCTGGAGGCCATATATGAAAGGCACTCATGG	7181
OY	7057	TAAATATGACCATGGGGAAGTCTCCGCGGCTCAGGCGCAATGCTTTTAACTCTTGAA	7116
Db	7182	TAAATATGACCATGGGGAAGTCTCCGCGGCTCAGGCGCAATGCTTTTAACTCTTGAA	7241
OY	7117	TGCCAGTGCACGCGTCCGCTGTATAGCCATPAACGCTGTCTGAAGGAGCGGATGACCA	7176
Db	7242	TGCCAGTGCACGCGTCCGCTGTATAGCCATPAACGCTGTCTGAAGGAGCGGATGACCA	7301
OY	7177	CACACTCACTCCGCAAGTGGCGGGGAAAGGCAAGGTCTCTGGCAGACCAGCAGACCG	7236
Db	7302	CACACTCACTCCGCAAGTGGCGGGGAAAGGCAAGGTCTCTGGCAGACCAGCAGACCG	7361
OY	7237	AAAAGCCAAGTCCCGGCGCCCGGGCTTGGCATCTGGGGACCGGCACTCTGTCTCTC	7296
Db	7362	AAAAGCCAAGTCCCGGCGCCCGGGCTTGGCATCTGGGGACCGGCACTCTGTCTCTC	7421
OY	7297	AGTGCACCTCGGAGGGGAACATGCAACCGCGCGGACCGGCTCAACACGCGGTGGGGAGA	7356
Db	7422	AGTGCACCTCGGAGGGGAACATGCAACCGCGCGGACCGGCTCAACACGCGGTGGGGAGA	7481
OY	7357	CAGGCGCTCGTCCGCAAGTTCACAGCCATTCCCTTAACACCCCTGATCTCATGCGGCTGCA	7416
Db	7482	CAGGCGCTCGTCCGCAAGTTCACAGCCATTCCCTTAACACCCCTGATCTCATGCGGCTGCA	7541
OY	7417	GGCGGGGTGCAATGGCTTCCCAACCCCAACCGGCGCTCCCGGGGACGCGGGCCCTCGC	7476
Db	7542	GGCGGGGTGCAATGGCTTCCCAACCCCAACCGGCGCTCCCGGGGACGCGGGCCCTCGC	7601
OY	7477	TGGCCCCCAACACGCTCTGGGACGAGGAGCCCAAGCCATCTCTCTGCTCGCATGACGAGAC	7536
Db	7602	TGGCCCCCAACACGCTCTGGGACGAGGAGCCCAAGCCATCTCTCTGCTCGCATGACGAGAC	7661
OY	7537	ACTCTCCGACGACGATGATCAAGAACAGGGCGGGGGGGGGGGCGGCTGTUERSPERLUG	7596
Db	7662	ACTCTCCGACGACGATGATCAAGAACAGGGCGGGGGGGGGGGCGGCTGTUERSPERLUG	7711
OY	7597	TCAAGTCCCAAGCAGCACAAGAAACGCGCTTCAGAGCGGGGCGCTGCGACTCCCC	7656
Db	7712	TCAAGTCCCAAGCAGCACAAGAAACGCGCTTCAGAGCGGGGCGCTGCGACTCCCC	7771
OY	7657	AACCAAGAAAGAGAGCCCTTGAATGTCCGCTGCGCTCATATCATATCTGCTCCAGAGCCG	7716
Db	7772	AACCAAGAAAGAGAGCCCTTGAATGTCCGCTGCGCTCATATCATATCTGCTCCAGAGCCG	7831
OY	7717	GCATCTTGTGCTGTCTAAAGCTTAACTAAGACTCCCGCCCGGGGCTGGGCTGTGTCAGA	7776
Db	7832	GCATCTTGTGCTGTCTAAAGCTTAACTAAGACTCCCGCCCGGGGCTGGGCTGTGTCAGA	7891
OY	7777	CCCTTACTCAGGGGATGTTTACTCTGTGTCTCGGAGGGGAGGGGAGGGGCGCGGAGAGGG	7836



Db	7892	CCTTA	CTCAGGGAGATGTTTACTGCTGCTCGGAGAAAGGAGGAAAGGGCCGGGGAGAGGG	7951
Qy	7837	GCACGG	CAGGCGTGTGGGACGCCACAACAAGCGGCGAGGCGGCCAGGGACCCAAAGCAG	7896
Db	7952	GCACGG	CAGGCGTGTGGGACGCCACAACAAGCGGCGAGGCGGCCAGGGACCCAAAGCAG	8011
Qy	7897	GATGAC	CCACGACCTCCACGCGCACTGGCTCCCGGAAATGCAATTTGGAAACCAAGCTTAA	7956
Db	8012	GATGAC	CCACGACCTCCACGCGCACTGGCTCCCGGAAATGCAATTTGGAAACCAAGCTTAA	8071
Qy	7957	CTGAGC	CTGCAAGCCCGCCGCGCCCTCCCTCCGCGCTCCCAATCCCGCTTAGCGCTTGACAG	8016
Db	8072	CTGAGC	CTGCAAGCCCGCCGCGCCCTCCCTCCGCGCTCCCAATCCCGCTTAGCGCTTGACAG	8131
Qy	8017	ATGGA	ACGAGGCCCCCTGTCAGCCCCCGACGTGCGCTGTTCCGGTCCCAACAGACTGCCCA	8076
Db	8132	ATGGA	ACGAGGCCCCCTGTCAGCCCCCGACGTGCGCTGTTCCGGTCCCAACAGACTGCCCA	8191
Qy	8077	GCCAA	CCGAGATTTGCTGGAACCAAGTCAGGCGCAGGTGGGCGGACAAAGGGCCAGGTGCG	8136
Db	8192	GCCAA	CCGAGATTTGCTGGAACCAAGTCAGGCGCAGGTGGGCGGACAAAGGGCCAGGTGCG	8251
Qy	8137	GCC	TGGGGGGGAAACGGATGCTCCGAGGACTGGACTGTTTTTTTTCACACATCGTGCCGAG	8196
Db	8252	GCC	TGGGGGGGAAACGGATGCTCCGAGGACTGGACTGTTTTTTTTCACACATCGTGCCGAG	8311
Qy	8197	CGGT	GGAGAAAGGAAAGCAGATGTAATGATGTGTGTTTACAGGGTATATTTTGTATAC	8256
Db	8312	CGGT	GGAGAAAGGAAAGCAGATGTAATGATGTGTGTTTACAGGGTATATTTTGTATAC	8371
Qy	8257	CTTCA	TGATATTAATTACAGATGTTTACGCAAGAAAGACTTACCCAGTATTACTGCTGC	8316
Db	8372	CTTCA	TGATATTAATTACAGATGTTTACGCAAGAAAGACTTACCCAGTATTACTGCTGC	8431
Qy	8317	TGTG	CTTTTGATCTCTGCTTACCGTTCAAGAGGCGGTGACAGGCCGACAGTCGGTACCC	8376
Db	8432	TGTG	CTTTTGATCTCTGCTTACCGTTCAAGAGGCGGTGACAGGCCGACAGTCGGTACCC	8491
Qy	8377	CATCA	CTCGCAGACCAAGGGGGGCGGAGCTGCTGTCACGCCCGCGCTGTCTCCCTC	8436
Db	8492	CATCA	CTCGCAGACCAAGGGGGGCGGAGCTGCTGTCACGCCCGCGCTGTCTCCCTC	8551
Qy	8437	CCTC	CCCTTCTTTGGGACAGATGATTCGATGCGTATCTGTGCGCCGTCATTGGCGCAGGG	8496
Db	8552	CCTC	CCCTTCTTTGGGACAGATGATTCGATGCGTATCTGTGCGCCGTCATTGGCGCAGGG	8611
Qy	8497	TGGT	GGTATTTCTGTCATTTTACACAGCTGCTCTAATTTAAAAAGCAATTATCTCCAAA	8556
Db	8612	TGGT	GGTATTTCTGTCATTTTACACAGCTGCTCTAATTTAAAAAGCAATTATCTCCAAA	8671
Qy	8557	AAAA	AAAAAAAAAAAAAA 8571	
Db	8672	AAAA	AAAAAAAAAAAAAA 8686	
RESULT 8				
ADL13811				
ID	ADL13811	standard; DNA; 8686 BP.		
XX	AC	ADL13811;		
XX	DT	06-MAY-2004 (first entry)		
DE	XX	Osteoarthritis-associated polymorphic nucleotide #343.		
XX	de;	gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;		
KW	joint	space narrowing; osteophyte development; joint pain;		
RW	osteoarthritis;	SNP; single nucleotide polymorphism.		
OS	Homo	sapiens.		
XX	PN	W02003054166-A2.		
XX				

PD	03-JUL-2003.
XX	
PF	19-DEC-2002; 2002WO-US041225.
XX	
PR	20-DEC-2001; 2001US-0342603P.
XX	
PA	(INCY-) INCYTE GENOMICS INC.
XX	
PI	Jones KA, Schafer A;
XX	
DR	WPI; 2003-559141/52.
XX	
PT	Determining susceptibility of an individual to joint space narrowing,
PT	osteophyte development and/or joint pain comprises identifying whether
PT	the individual has at least one polymorphism in a polynucleotide encoding
PT	a protein.
XX	
PS	Disclosure; SEQ ID NO 343; 297bp; English.
XX	
CC	The invention relates to a method of determining susceptibility of an
CC	individual to joint space narrowing and/or osteophyte development and/or
CC	joint pain comprising identifying whether the individual has at least one
CC	polymorphism in a polynucleotide encoding at least one of the protein
CC	listed in the specification. The methods, composition and agent are
CC	useful for modulating the susceptibility of an individual to joint space
CC	narrowing and/or osteophyte development and/or joint pain that is
CC	associated with a disease, preferably osteoarthritis. The cell line and
CC	the non-human animal are useful for screening for an agent for diagnosing
CC	an individual having susceptibility to joint space narrowing and/or
CC	osteophyte development and/or joint pain. This sequence corresponds to
CC	the polynucleotide encoding a protein listed in the specification. (Note:
CC	The sequence data for this patent did not form part of the printed
CC	specification but was obtained in electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences).
XX	
XX	Sequence 8686 BP; 1881 A; 3069 C; 2562 G; 1174 T; 0 U; 0 Other;
SQ	
Query Match	97.2%; Score 8324.6; DB 10; Length 8686;
Best Local Similarity	98.6%; Pred. No. 0;
Matches 8472; Conservative	1; Mismatches 34; Indels 88; Gaps 4
OY	1 CATGTGGGCTCCACACAGCTTTGTGGCAGACGTGAGGGGCACCTAGGCCCTTACC 60
DB	156 CATGTGGGCTCCACACAGCCTGTGTGGCACACAGCGTAGGGCCCACTGAGCCCCCTTACC 215
OY	61 GCCCCACAGCCTTTCTTAACCAATGCGCCCGACACACAGACGTGGGCTCT 120
DB	216 GCCCCACAGCCTTTCTTAACCAATGCGCATATGCGCCCGACACACAGACGTGGGCTCT 275
OY	121 GAAGTACAGACACACTCCCGCATGACCTCCACACTGTTCGCGCGGCTCCATATCCA 180
DB	276 GAAGTACAGACACACTCCCGCATGACCTCCACACTGTTCGCGGCTCCATATCCA 335
OY	181 GCCCCAGCGCGAGAGCCCTCTCTGTGTGAAGTTCAGGCCCGGAAATGAACGCTCCA 240
DB	336 GCCCCAGCGCGAGAGCCCTCTCTGTGTGAAGTTCAGGCCCGGAAATGAACGCTCCA 395
OY	241 GGAGTCCACCTGGGGCCAGATGCCATCACTACCTGCGCCGAGCGGGGAATCAGAAT 300
DB	396 GGAGTCCACCTGGGGCCAGATGCCATCACTACCTGCGCCGAGCGGGGAATCAGAAT 455
OY	301 GAAGTCAATTAAAAGCAAGCCCTCGGCTAGAGCTGTGCTGACCCCTCTGTCGACC 360
DB	456 GAAGTCAATTAAAAGCAAGCCCTCGGCTAGAGCTGTGCTGACCCCTCTGTCGACC 515
OY	361 GTCAACCTCTGCTGGCCACGGGCGACCTTGCGGATCTGAAGACTCACCAAGAACCTGAG 420
DB	516 GTCAACCTCTGCTGGCCACGGGCGACCTTGCGGATCTGAAGAACCTCACCAAGAACCTGAG 575
OY	421 CCTAAGCGGGCAAGCTGGAACCGGTGTCCCCCAAGCCCCCGACACATGAACCTGAGCT 480
DB	576 CCTAAGCGGGCAAGCTGGAACCGGTGTCCCCCAAGCCCCCGACACATGAACCTGAGCT 635

481 GGAGCTGTGTCGCCCAACGGCTGTCCAAGAGAGCTGATCCAGAACATGGAACCGCGTGA 540  
636 GGAGCTGTGTGCGGCCCAACGGCTGTCCAAGAGAGAGCTGATCCAGAACATGGAACCGCGTGA 695  
541 CCGAGAGATCAACCATGTGTAGAGAGAGAGATCTTTAAAGCTGAGAGAGAGAGAGAGAGCT 600  
696 CCGAGAGATCAACCATGTGTAGAGAGAGAGATCTTTAAAGCTGAGAGAGAGAGAGAGAGCT 755  
601 GGAG 660  
756 GGAG 815  
661 GTTCGAAGACCCGAGAGCTGTGTGAGATCATCTACGAGAGAGAGAGAGAGAGAGAGAGAG 720  
816 GTTCGAAGACCCGAGAGCTGTGTGAGATCATCTACGAGAGAGAGAGAGAGAGAGAGAGAG 875  
721 TGCACATCGGATTTCTGAGAGAGCTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780  
876 TGCACATCGGATTTCTGAGAGAGCTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 935  
781 CTCGCCACACCCGAGAGATCATCTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840  
936 CTCGCCACACCCGAGAGATCATCTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 995  
841 AATCTTGTACTTCAAGAGAGAGAGATCATCGCTCGAGAGAGAGAGAGAGAGAGAGAGAG 900  
996 AATCTTGTACTTCAAGAGAGAGAGATCATCGCTCGAGAGAGAGAGAGAGAGAGAGAGAG 1055  
901 GCGCTATGACACAGCTCATGAGAGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960  
1056 GCGCTATGACACAGCTCATGAGAGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1115  
961 GCGCCCGCGCGGCCCAAG 1020  
1116 CCGGCTGAGACAGCTCATGAGAGAGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1175  
1021 CCGGAG 1080  
1176 CCGGAG 1232  
1081 GCTGTTCATGTGTGAG 1140  
1233 GCTGTTCATGTGTGAG 1292  
1141 AAG 1200  
1293 AAG 1352  
1201 CGAGCTGACAG 1260  
1353 CGAGCTGACAG 1412  
1261 GAAAGGTGTCAAAAGACCGGCAAGGTCAATGAAATGTGAGAGAGAGAGAGAGAGAGAGAG 1320  
1413 GAAAGGTGTCAAAAGACCGGCAAGGTCAATGAAATGTGAGAGAGAGAGAGAGAGAGAGAG 1472  
1321 CCGGAG 1380  
1473 CCGGAG 1532  
1381 GAAAGACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1440  
1533 GAAAGACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1592  
1441 GAGGCTGTGTGAG 1500  
1593 GAGGCTGTGTGAG 1652  
1501 GCAAG 1560  
1653 GCAAG 1712  
1561 AGATGAG 1620

1713 AGATGAG 1772  
1621 CGAACAG 1680  
1773 CGAACAG 1832  
1681 GAAAGAGCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1740  
1833 GAAAGAGCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1892  
1741 CCGCATCACCCGCTCAATGAGCTTAATGAGGCAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1800  
1893 CCGCATCACCCGCTCAATGAGCTTAATGAGGCAACAGAGAGAGAGAGAGAGAGAGAGAGAG 1952  
1801 GAGCGCGAGAGCTGTGTGAG 1860  
1953 GAGCGCGAGAGCTGTGTGAG 2012  
1861 GGAAACAGGCAAGAGAGAGTCTCTGTGAGACACCGGCGGCACTGTGTGTGAGAGAGAGAG 1920  
2013 GGAAACAGGCAAGAGAGAGTCTCTGTGAGACACCGGCGGCACTGTGTGTGAGAGAGAGAG 2072  
1921 GGTGTGCTCCAGAGCTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1980  
2073 GGTGTGCTCCAGAGCTGTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2132  
1981 GAACTGTGATGAGATTTTGTGAG 2040  
2133 GAACTGTGATGAGATTTTGTGAG 2192  
2041 GAGGAG 2100  
2193 GAGGAG 2252  
2101 GAGGAG 2160  
2253 GAGGAG 2212  
2161 GGTGTGAGCTTTTCAATGCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220  
2213 GGTGTGAGCTTTTCAATGCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2272  
2221 CACTGTCAACACAGGCTTCAAG 2280  
2322 CACTGTCAACACAGGCTTCAAG 2381  
2281 GGAACAGAGGAGAGATGTGAG 2340  
2382 GGAACAGAGGAGAGATGTGAG 2441  
2341 AGGCGCACCCACCCCAACAG 2400  
2442 AGGCGCACCCACCCCAACAG 2501  
2401 TGAAGGCAACGAG 2460  
2502 TGAAGGCAACGAG 2561  
2562 TGTGTGTCCCAAG 2621  
2521 GGAAG 2580  
2622 GGAAG 2681  
2581 GCCCGTCAAG 2640  
2682 GCCCGTCAAG 2741  
2641 GGAAGGCGCTGTGAG 2700

Db 2742 GAGGCCGCTGAGGCCACGGCCGAGAGGGCGCTCAAGGCAAGAAAGAGAGGGCGGGAG 2801  
Qy 2701 CGGCAAGGGCCACCACTGCCAAGAGCTCGGGCGCCCCCAGAGCAGGCACTCCAGTGTAC 2760  
Db 2802 CGGCAAGGGCCACCAAGAGGCTCGGGCGCCCCCAGAGCAGGCACTCCAGTGTCCAC 2861  
Qy 2761 CTGCAGTGCAGACGAGGTGATGAGGCCGAGGGCGGCGCAAGAAACCGGCTGTGCCCC 2820  
Db 2862 CTGCAGTGCAGACGAGGTGATGAGGCCGAGGGCGGCGCAAGAAACCGGCTGTGCCCC 2921  
Qy 2821 AAGGCCGAGCCTCTCAACCCCGACTGGGCGACCCCGGGGCAATGCTCAACCCAGAGGC 2880  
Db 2922 AAGGCCGAGCCTCTCAACCCCGACTGGGCGACCCCGGGGCAATGCTCAACCCAGAGGC 2981  
Qy 2881 ACTGCACTGAAGCAGCTGGAAGCAGGAGCGGCTGCATCCCCCATCCAGTCACTCA 2940  
Db 2982 ACTGCACTGAAGCAGCTGGAAGCAGGAGCGGCTGCATCCCCCATCCAGTCACTCA 3041  
Qy 2941 AGTCCATGAGCCCCCGGAGAGAGCGAGCTCCACCAAGCCAGTCCCCCAGCCCCCACC 3000  
Db 3042 AGTCCATGAGCCCCCGGAGAGAGCGAGCTCCACCAAGCCAGTCCCCCAGCCCCCACC 3101  
Qy 3001 GCCACCGCAAAACTGCAAGCCGAGAGCGAGCCCTCGACGAGCCTGGCAGAGGCCCGG 3060  
Db 3102 GCCACCGCAAAACTGCAAGCCGAGAGCGAGCCCTCGACGAGCCTGGCAGAGGCCCGG 3161  
Qy 3061 GGGCAAGACAGAGAGCCCGGCAACCCCGCGCGCAAGAGAG----- 3100  
Db 3162 GGGCAAGACAGAGAGCCCGGCAACCCCGCGCGCAAGAGAGAGAGCTGTGTCTT 3221  
Qy 3101 ----GCTTTCGAGCGGAGGCCCAAGAGCTGCTGGGAGCCCCCTTGTGTGATTTCCGG 3156  
Db 3222 CCACACCTTCGAGCGGAGGCCCAAGAGCTGCTGGGAGCCCCCTTGTGTGATTTCCGG 3281  
Qy 3157 CCTGCGCTTCCCGGAGCGCGCGCGGAGGTGATCAAGGCTCCCGGCATGCGCCGAGACCC 3216  
Db 3282 CCTGCGCTTCCCGGAGCGCGCGCGGAGGTGATCAAGGCTCCCGGCATGCGCCGAGACCC 3341  
Qy 3217 CTCAAGCTTCTCTCAAGCTTCACTGCTCACTGCTCACTGCTGAGGCTTCCATGACATGAC 3276  
Db 3342 CTCAAGCTTCTCTCAAGCTTCACTGCTCACTGCTCACTGCTGAGGCTTCCATGACATGAC 3401  
Qy 3277 CGGCGCGCTTCTGCG 3336  
Db 3402 CGGCGCGCTTCTGCG 3461  
Qy 3337 CAAGCACCCGAGCGCTCTGAGAGGCAATAGGTGCAATCTCCAGAGATGTCGCTCCA 3396  
Db 3462 CAAGCACCCGAGCGCTCTGAGAGGCAATAGGTGCAATCTCCAGAGATGTCGCTCCA 3521  
Qy 3397 GCTCCAGCTCCCGTACTGAGAGCATGCCAAGGCCCGGCTGAGCGCTGTCAACATGGGCT 3456  
Db 3522 GCTCCAGCTCCCGTACTGAGAGCATGCCAAGGCCCGGCTGAGCGCTGTCAACATGGGCT 3581  
Qy 3457 GCGCTTGCCTATGAGACCCCAAAAAGCTGGCACCTTCAAGCGAGTGAACAGAGCAGCT 3516  
Db 3582 GCGCTTGCCTATGAGACCCCAAAAAGCTGGCACCTTCAAGCGAGTGAACAGAGCAGCT 3641  
Qy 3517 GTCCCAAGGGGCGAGGCTGGGCGACCGGAGAGCTGGGGGTGGCCACAGGCCAGAGAGGC 3576  
Db 3642 GTCCCAAGGGGCGAGGCTGGGCGACCGGAGAGCTGGGGGTGGCCACAGGCCAGAGAGGC 3701  
Qy 3577 GTCCGTGCTGAGAGGAGAGCTCTGGGCTCAAGTTCCGGGCGGAGACATCAACAAAGGCT 3636  
Db 3702 GTCCGTGCTGAGAGGAGAGAGCTCTGGGCTCAAGTTCCGGGCGGAAACATCAACAAAGGCT 3761  
Qy 3637 TCCCAAGCAACAGGGTGCCTTGGAGCAAGCGCATCAATACCGGCTTCATCAACCCACGG 3696  
Db 3762 TCCCAAGCAACAGGGTGCCTTGGAGCAAGCGCATCAATACCGGCTTCATCAACCCACGG 3821  
Qy 3697 CACGCAAGCTGACGTCGTATCAAGGGGCAACATCAACAGAGATCAACGAGGAGAGAGCC 3756  
Db 3822 CACGCAAGCTGACGTCGTATCAAGGGGCAACATCAACAGAGATCAACGAGGAGAGAGCC 3881

Qy 3757 GAGTCGCTTGAACCGCGGCGCGGAGAGACAGCTGTGCCAAGGGCCAGCTATCTACGAAGG 3816  
Db 3882 GAGTCGCTTGAACCGCGGCGCGGAGAGACAGCTGTGCCAAGGGCCAGCTATCTACGAAGG 3941  
Qy 3817 CAAGAAGGGCCACGCTTGTCTTATGAGGGTGGCATGTCTGTGACCCAGTCTCCAGGA 3876  
Db 3942 CAAGAAGGGCCACGCTTGTCTTATGAGGGTGGCATGTCTGTGACCCAGTCTCCAGGA 4001  
Qy 3877 GGAAGGCAAGACAGCTCAGGACCCCGCATGAGACGGCGGCGCCCAAGCGCAGCTATGA 3936  
Db 4002 GGAAGGCAAGACAGCTCAGGACCCCGCATGAGACGGCGGCGCCCAAGCGCAGCTATGA 4061  
Qy 3937 CATGATGAGGGGCGCGTGGGAGAGCCATCTCTCAAGCAGCATGAGGTCTCATGGG 3996  
Db 4062 CATGATGAGGGGCGCGTGGGAGAGCCATCTCTCAAGCAGCATGAGGTCTCATGGG 4121  
Qy 3997 CGGTGCTATCCCGCGGAGAGCCACAGGCCCGCCACACTTCAAAAGGAGACACATCCG 4056  
Db 4122 CGGTGCTATCCCGCGGAGAGCCACAGGCCCGCCACACTTCAAAAGGAGACACATCCG 4181  
Qy 4057 CGGTGCTATCAACAAGGGATCCCTCGGTCTTACGTGAGGAGCAGAGAGACTTACTGGG 4116  
Db 4182 CGGTGCTATCAACAAGGGATCCCTCGGTCTTACGTGAGGAGCAGAGAGACTTACTGGG 4241  
Qy 4117 TCGGAGGGCCAAAGCTCTTAAAGCGGAGGGCAAGCTTCGCCCCCAAGCGCTCAAGGGA 4176  
Db 4242 TCGGAGGGCCAAAGCTCTTAAAGCGGAGGGCAAGCTTCGCCCCCAAGCGCTCAAGGGA 4301  
Qy 4177 CTTGACCGAGGCTTACAAAGACGAGGCCCTTGGGCGCCCTTGAAGCTGAAGCGGCGCATGA 4236  
Db 4302 CTTGACCGAGGCTTACAAAGACGAGGCCCTTGGGCGCCCTTGAAGCTGAAGCGGCGCATGA 4361  
Qy 4237 GGGCGCTGGTGGCAAGGTAAGAGAGAGCGGAGCGGCTCCATCCATGAGATCCCGCGAGGA 4296  
Db 4362 GGGCGCTGGTGGCAAGGTAAGAGAGAGCGGAGCGGCTCCATCCATGAGATCCCGCGAGGA 4421  
Qy 4297 GCTGCGGACACAGCCCGAGCTGCGCTTGGGCCCGCGGCGGCTCAAGAGGGCTCCATCAC 4356  
Db 4422 GCTGCGGACACAGCCCGAGCTGCGCTTGGGCCCGCGGCGGCTCAAGAGGGCTCCATCAC 4481  
Qy 4357 GCAAGGACACCCCGCTCAAGTACAGACCGGCGGCTTCCACACTGGCTCCAAAAGACAGA 4416  
Db 4482 GCAAGGACACCCCGCTCAAGTACAGACCGGCGGCTTCCACACTGGCTCCAAAAGACAGA 4541  
Qy 4417 GGTACGCTCCGTCATGAGCAGGCCCGGAGCGTGTCCACCGGAGCACCGGCTGAGTGT 4476  
Db 4542 GGTACGCTCCGTCATGAGCAGGCCCGGAGCGTGTGTCCACCGGAGCACCGGCTGAGTGT 4601  
Qy 4477 GATGGCCGAGCGCCCGGAGCTTGAACGTGCTGTACAGAGAGAGCTTGAAGAGCGGCGC 4536  
Db 4602 GATGGCCGAGCGCCCGGAGCTTGAACGTGCTGTACAGAGAGAGCTTGAAGAGCGGCGC 4661  
Qy 4537 AGGAGCCGCGAGAGCTGGGGGGGCTCCATTGGCGGCGGCGCCCGGTGATTTGTGCTGA 4596  
Db 4662 AGGAGCCGCGAGAGCTGGGGGGGCTCCATTGGCGGCGGCGCCCGGTGATTTGTGCTGA 4721  
Qy 4597 GGTGGGTAAAGCGCGGAGAGGCCCTGACCTTATGAGGACCAAGGGGAGCACTTTTGGCGG 4656  
Db 4722 GGTGGGTAAAGCGCGGAGAGGCCCTTGAACCTTATGAGGACCAAGGGGAGCACTTTTGGCGG 4781  
Qy 4657 CCACCTCCACAGAGTTTGCCTGTGACCATGCGGAGGCCACGCGCGCGCTTGCAGAGAGG 4716  
Db 4782 CCACCTCCACAGAGTTTGCCTGTGACCATGCGGAGGCCACGCGCGCGCTTGCAGAGAGG 4841  
Qy 4717 CAGCCTTTGTTCAGCAAGGACATCCCAGAGCCGAAGCTGACGTCGCTCGTGAAT 4776  
Db 4842 CAGCCTTTGTTCAGCAAGGACATCCCAGAGCCGAAGCTGACGTCGCTCGTGAAT 4901  
Qy 4777 GCGCAAGTCCCGGACAGGACCGTGGCCGAGACCAACCAACCCATCTGGCCCTATGA 4836  
Db 4902 GCGCAAGTCCCGGACAGGACCGTGGCCGAGACCAACCAACCCATCTGGCCCTATGA 4961

OY	4837	GCACTTGTTGGGGCGTGAAGTGGGTGA	CTGTATGTGAGCCACATCCCCGTGGCTT	4838
Db	4962	GACCTGCTTGGGGCGTGAAGTGGGTGA	CTGTATGTGAGCCACATCCCCGTGGCTT	5021
OY	4897	GCAACCCACCTCAATACCCCGGGCATCC	CTTGAGCGAGCGGTGGCTACTACCTGCC	4956
Db	5022	GGAACCCACCTTCATACCCCGGGCATCC	CTTGAGCGAGCGGTGGCTACTACCTGCC	5081
OY	4957	CCGACACTGGCCCCCAACCCCACTTAC	CCCGCACTGTATACCCCACTTATCCGGG	5016
Db	5082	CCGACACTGGCCCCCAACCCCACTTAC	CCCGCACTGTATATCCCACTTATCCGGG	5141
OY	5017	CTACCCCGACACGGGGGGGCTTGAAGA	ACCGGACATCATCATATGATCATCACTTC	5076
Db	5142	CTACCCCGACACGGGGGGGCTTGAAGA	ACCGGACATCATCATATGATCATCACTTC	5201
OY	5077	GCAGCAGATGCAACAACAACGGCCACA	CCGGCCATTTGGCCAGAGCTGATATGTGAGGGG	5136
Db	5202	GCAGCAGATGCAACAACAACGGCCACA	CCGGCCATTTGGCCAGAGCTGATATGTGAGGGG	5261
OY	5137	CCTCTCGCCCCGCGAGTCTCTGCTGGCA	CTCAACTACGCTGGGGTCCCGAGGATCAT	5196
Db	5262	CCTCTCGCCCCGCGAGTCTCTGCTGGCA	CTCAACTACGCTGGGGTCCCGAGGATCAT	5321
OY	5197	CGACTGTGTCCAAATGTGCCACATCTG	CTGTGCTGTGCTCCGCCGACACAGCAACCCACG	5256
Db	5322	CGACTGTGTCCAAATGTGCCACATCTG	CTGTGCTGTGCTCCGCCGACACAGCAACCCACG	5381
OY	5257	CACCGCCATNGACCGCGCTTGGCTACT	CCCTCCCAACCGCGCCCAAGCCCTTACAGACGGCA	5316
Db	5382	CACCGCCATNGACCGCGCTTGGCTACT	CCCTCCCAACCGCGCCCAAGCCCTTACAGACGGCA	5441
OY	5317	CAGCAGCTCTCCCACTCTCCCCAGAGAT	TCACAACATTTGACAAACCAACCAACACGCTC	5376
Db	5442	CAGCAGCTCTCCCACTCTCCCCAGAGAT	TCACAACATTTGACAAACCAACCAACACGCTC	5501
OY	5377	CTCGTCCGAGCGGGAGCGAGACCGGG	AATCGAAGCGGGAATCGAAGCGGGACCGGGATCGGGAGCGGGAAAA	5436
Db	5502	CTCGTCCGAGCGGGAGCGAGACCGGG	AATCGAAGCGGGAATCGAAGCGGGACCGGGATCGGGAGCGGGAAAA	5561
OY	5437	GTCATCTCTTCACTGTCCACACAGACG	GTGTGAAGACACCCATCTGTGAAGACCTTGTATACGA	5486
Db	5562	GTCATCTCTTCACTGTCTCACACAGACG	GTGTGAAGACACCCATCTGTGAAGACCTTGTATACGA	5621
OY	5497	GCAGAGCAGCGGCAGACCGGCAGACG	GGCGGGGTGGGGGACAGCAGACCGGCCGC	5556
Db	5622	GCAGAGCAGCGGCAGACCGGCAGACG	GGCGGGGTGGGGGACAGCAGACCGGCCGC	5681
OY	5557	CTCCCACTCCCATGCCCCACAGCACT	CGCCCATCTCCCCCTGGAACCAAGATGCCCTTCA	5616
Db	5682	CTCCCACTCCCATGCCCCACAGCACT	CGCCCATCTCCCCCTGGAACCAAGATGCCCTTCA	5741
OY	5617	GCAAGACCCCATGTGTCTTACACAAC	ACAGGCATGAAGGTATCATCACCGCTGTGAAGCC	5676
Db	5742	GCAAGACCCCATGTGTCTTACACAAC	ACAGGCATGAAGGTATCATCACCGCTGTGAAGCC	5801
OY	5677	CAGCAAGGCCACGGTCTGAAGGTTCAC	CTCCACCTCCACCCGTTGGCCCAAGCTGGCAC	5736
Db	5802	CAGCAAGGCCACGGTCTGAAGGTTCAC	CTCCACCTCCACCCGTTGGCCCAAGCTGGCAC	5861
OY	5737	ATTCCCACTTGACACCACTGCCCCACT	GTGGCGGCAACCTCTGATGGGGTCTAACCTTACCT	5796
Db	5862	ATTCCCACTTGACACCACTGCCCCACT	GTGGCGGCAACCTCTGATGGGGTCTAACCTTACCT	5921
OY	5797	CATGAGCCCCGTCTTGTCTGCCAGAG	AGGCCCCCGGGTCCGCCAGACCGGCCCCG	5856
Db	5922	CATGAGCCCCGTCTTGTCTGCCAGAG	AGGCCCCCGGGTCCGCCAGACCGGCCCCG	5981
OY	5857	AGCAGACACCGGCGCATTTGCTTCTG	CCAAAGCCCCCAGCCCGGCTCCGGGGCTGGAAGCCGC	5916
Db	5982	AGCAGACACCGGCGCATTTGCTTCTG	CCAAAGCCCCCAGCCCGGCTCCGGGGCTGGAAGCCGC	6041
OY	5917	CTCCTCCCCCAGCAAGGGCTGGAAGCC	CCCGGCCCTTATGTGCTCTCTGTCTCTGGCACGC	5976

Db	6042	CTCTCTCCCCAGCAAGGGCTGAGAGCCCGGCCCTTAGTGCTCTGTCTTGGCCAGGC	6101
Qy	5977	CACCATGCCCCGCAACCCCTGCGAAGAACTTGCACTTCACCGCCAGCCCGGACCCGCC	6036
Db	6102	CACCATGGCCCCGCAACCCCTGCGAAGAACTTGCACTTCACCGCCAGCCCGGACCCGCC	6161
Qy	6037	GGCGCCACTGTCCTCGGCTCTTGGAACCGGCACCGGGAAAAAGCTCAAGTAACCTTTTC	6096
Db	6162	GGCGCCACTGTCCTCGGCTCTTGGAACCGGCACCGGGAAAAAGCTCAAGTAACCTTTTC	6221
Qy	6097	CATCCAGGAATGGAATCTCGGTTCTTGGGTTCACAGCGGAGAGCTACAGCCCCGAAG	6156
Db	6222	CATCCAGGAATGGAATCTCGGTTCTTGGGTTCACAGCGGAGAGCTACAGCCCCGAAG	6281
Qy	6157	GGTGAAGCCCTCAAGCCCTGTGACTCAACCACTGTGACCCACGACAAAGGGCTTCCCA	6214
Db	6282	GGTGAAGCCCTCAAGCCCTGTGACTCAACCACTGTGACCCACGACAAAGGGCTTCCCA	6341
Qy	6217	GCACCTGGAGAGCTCGACAAAGAGCCACTTGGAAGGGAGCTGCGGCCAAGCCACAG	6276
Db	6342	GCACCTGGAGAGCTCGACAAAGAGCCACTTGGAAGGGAGCTGCGGCCAAGCCACAG	6401
Qy	6277	CCCCGTGAAGTTTGGCGGGGAGGGCGGCCACTGCCACAACCTGGGGCGCTGCGTAG	6336
Db	6402	CCCCGTGAAGTTTGGCGGGGAGGGCGGCCACTGCCACAACCTGGGGCGCTGCGTAG	6461
Qy	6337	CCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGGCCAGGGGTCAAAAGTCAACAGCGGGT	6396
Db	6462	CCAGCCCTCGTCCAGCCCGCTGCTCCAGACCGGCCAGGGGTCAAAAGTCAACAGCGGGT	6521
Qy	6397	GGTCAACCTGCGCCAGCAATCACTAGTAGGTATCAACAAGACTAACCCGGCACACCC	6456
Db	6522	GGTCAACCTGCGCCAGCAATCACTAGTAGGTATCAACAAGACTAACCCGGCACACCC	6581
Qy	6457	ACAGCAGTCAAGGGCAACCTTGCGGCCGCCCTCTTACTCTTCCCTGGGGGCACTGGCC	6516
Db	6582	ACAGCAGTCAAGGGCAACCTTGCGGCCGCCCTCTTACTCTTCCCTGGGGGCACTGGCC	6641
Qy	6517	CGTCTTGGAACTTCGCGCGGCCCAACCACTGACCTTCACTCCGCGCCCGGACATGTC	6576
Db	6642	CGTCTTGGAACTTCGCGCGGCCCAACCACTGACCTTCACTCCGCGCCCGGACATGTC	6701
Qy	6577	CCCGGCCCGTGGCTCCCCCAACAGCGAAGGGGCAAGAGGTCTTCAGAGCCAAACAAGC	6636
Db	6702	CCCGGCCCGTGGCTCCCCCAACAGCGAAGGGGCAAGAGGTCTTCAGAGCCAAACAAGC	6761
Qy	6637	GTCCGTCTTGGGTGTGTGTGAGACGGTATTGAACCTGTGTCTCCCACTGGAGGGCATGAC	6696
Db	6762	GTCCGTCTTGGGTGTGTGTGAGACGGTATTGAACCTGTGTCTCCCACTGGAGGGCATGAC	6821
Qy	6697	GGAGCCAGGGGCACTCCGGAGTGCTGTGTAACCGCTCTGTACCGGGATGCGGAACAAGC	6756
Db	6822	GGAGCCAGGGGCACTCCGGAGTGCTGTGTAACCGCTCTGTACCGGGATGCGGAACAAGC	6881
Qy	6757	GGAAGCCAGAGATGAGGCTTCAAGTCTCCAGGCAACAACAGCCAGCGCCAGCTTCTT	6816
Db	6882	GGAAGCCAGAGATGAGGCTTCAAGTCTCCAGGCAACAACAGCCAGCGCCAGCTTCTT	6941
Qy	6817	CAGCAAGCTGACGAGAGCAATTCGCCCATGTGTCAAGTCCAGAAAGCAGAGATCAACA	6876
Db	6942	CAGCAAGCTGACGAGAGCAATTCGCCCATGTGTCAAGTCCAGAAAGCAGAGATCAACA	7001
Qy	6877	GAACTGAACAACCAACAGCGGAATGAGCTGAATCAATTCAGCCAGCCTGGGAAGGA	6936
Db	7002	GAACTGAACAACCAACAGCGGAATGAGCTGAATCAATTCAGCCAGCCTGGGAAGGA	7061
Qy	6937	GATCTTCAATATGCCCCCATCAACCGGAACAGGCTTATGACTTATGAAGCCAGCGGT	6996
Db	7062	GATCTTCAATATGCCCCCATCAACCGGAACAGGCTTATGACTTATGAAGCCAGCGGT	7121
Qy	6997	GCAGGAACATGCCAGCAACAATGGGGCTTGGAGGCCAATTAATGAAGGCACTCATGGG	7056

Db	7122	GCAGGAACATGCAAGCACCAACATGTTGGGCTGGAGGCCATTATTTGAAGAGCACTCATGCG	7121
OY	7057	TAAATATGACAAGATGGGAAGAGTCCCGGCGCTCAGGCGCAATGCTTTAAACCTCTGAA	7116
Db	7182	TAAATATATGACAAGATGGGAAGAGTCCCGGCGCTCAGGCGCAATGCTTTAAACCTCTGAA	7241
OY	7117	TGGCAGTGCACAGCCCGCGGCTGCTATGCGCATMAACGCTGCTGACGGAGGAGACCA	7176
Db	7242	TGGCAGTGCACAGCCCGCGGCTGCTATGCGCATMAACGCTGCTGACGGAGGAGACCA	7301
OY	7117	CACACTCACTCGCAGGATGGCGGCGGGAAGGACAAAGTCTCTGGCACAACCAGACGCG	7236
Db	7302	CACACTCACTCGCAGGATGGCGGCGGGAAGGACAAAGTCTCTGGCACAACCAGACGCG	7361
OY	7237	AAAAGCCAGTCCCGGCGCCCGGCGCTGGCATCTGGGGACCGGCCACCTCTGTCTCCTC	7296
Db	7362	AAAAGCCAGTCCCGGCGCCCGGCGCTGGCATCTGGGGACCGGCCACCTCTGTCTCCTC	7421
OY	7287	AGTGCACCTCGAGGGGAGACTGCAACCGCGGACCGCGCTGACCAACCGCGTGTGGAGGA	7356
Db	7422	AGTGCACCTCGAGGGGAGACTGCAACCGCGGACCGCGCTGACCAACCGCGTGTGGAGGA	7481
OY	7357	CAGGCGCTCTGTCGCGAGGTTCCACGCGCATTCGCCCTCAACCCCGCTGATCATGCGCTGCA	7416
Db	7482	CAGGCGCTCTGTCGCGAGGTTCCACGCGCATTCGCCCTCAACCCCGCTGATCATGCGCTGCA	7541
OY	7417	GCGGGGTGATATGCTTCCCAACCCGCCACCGGGCTCCCGCGGACGGGACCCCTTCGC	7476
Db	7542	GCGGGGTGATATGCTTCCCAACCCGCCACCGGGCTCCCGCGGACGGGACCCCTTCGC	7601
OY	7417	TGGCCCCCAACCGCTGGGAGACGGAGACCCCAAGCACTGCTCTGCTGCGAGTACGAGAC	7536
Db	7602	TGGCCCCCAACCGCTGGGAGACGGAGACCCCAAGCACTGCTCTGCTGCGAGTACGAGAC	7661
OY	7537	ACTCTCCGACAGCGAGTACTCAGAAACGAGCGGGGGGGCGGGCGGTBTUBERPERLUG	7596
Db	7662	ACTCTCCGACAGCGAGTACTCAGAAACGAGCGGGGGGGGGCGGGCGGTBTUBERPERLUG	7711
OY	7597	TCAGGTCCCAAGACCAACGAAACGAGCTTTCAGAGAGCGGGCGGCTGCAGACTCCCGC	7656
Db	7712	TCAGGTCCCAAGACCAACGAAACGAGCTTTCAGAGAGCGGGCGGCTGCAGACTCCCGC	7711
OY	7657	AAACCAAGAAAGAGGCCCTTGAGTCCGCTTGCGCTCCATCCATCTGTCCGTCCAGAGCCG	7716
Db	7712	AAACCAAGAAAGAGGCCCTTGAGTCCGCTTGCGCTCCATCCATCTGTCCGTCCAGAGCCG	7831
OY	7717	GCATCTTTCGCTGTAAAGCCTTAACCTAAGACTCCCGGCCCGGCTGAGCCCTGACGA	7776
Db	7832	GCATCTTTCGCTGTGTAAAGCCTTAACCTAAGACTCCCGGCCCGGCTGAGCCCTGACGA	7891
OY	7777	CCTTACTCAGGGAGTGTTAACCTGTGTGTCGGGAAAGGAGGGGACCGGAGAGGGG	7836
Db	7892	CCTTACTCAGGGAGTGTTAACCTGTGTGTCGGGAAAGGAGGGGACCGGAGAGGGG	7951
OY	7837	GCAAGGACAGGGGTGTGTGACGCCACACACAGGCGGCTCAGGGCGGCAAGGACCCAAAGCAG	7896
Db	7952	GCAAGGACAGGGGTGTGTGACGCCACACACAGGCGGCTCAGGGCGGCAAGGACCCAAAGCAG	8011
OY	7897	GATGACCAAGCACTCCCAAGGCACTGCGCTCCCGCGGAATGCAATTTTGGAACCAAAAGCTAAA	7956
Db	8012	GATGACCAAGCACTCCCAAGGCACTGCGCTCCCGCGGAATGCAATTTTGGAACCAAAAGCTAAA	8071
OY	7957	CTGAGCTTCAGAGCCCGCGGCTTCCTCCGCTCCCATCCCGCTTAAAGCTTGGACAG	8016
Db	8072	CTGAGCTTCAGAGCCCGCGGCTTCCTCCGCTCCCATCCCGCTTAAAGCTTGGACAG	8131
OY	8017	ATGGAACGAGGCGCTGTCTCAGGCCCCCAAGTGCGCTGTTCGCGTCCCAACAGACTGGCCCA	8076
Db	8132	ATGGAACGAGGCGCTGTCTCAGGCCCCCAAGTGCGCTGTTCGCGTCCCAACAGACTGGCCCA	8191
OY	8077	GCCAAACGAGATTGGCTGGAAACCAAGTCAAGGCAAGTGGCGGACAAAGGGGCGAGTGGC	8136
Db	8192	GCCAAACGAGATTGGCTGGAAACCAAGTCAAGGCAAGTGGCGGACAAAGGGGCGAGTGGC	8251

OY	8137	GCCCTGGGCGGAACGGATGCTCCAGAGACTGGA	CTGTTTGTTCACA	CATCGTGC	CGCAG	8196
Db	8252	GCCCTGGGCGGAACGGATGCTCCAGAGACTGGA	CTGTTTGTTCACA	CATCGTGC	CGCAG	8311
OY	8197	CGGTGGGAAGAAAGCGAGATGTAATA	TGTTGTTTACAGG	GTATATTTTGATAC		8256
Db	8312	CGGTGGGAAGAAAGCGAGATGTAATA	TGTTGTTTACAGG	GTATATTTTGATAC		8371
OY	8257	CTTCATGAATTAAATTCAGATGTTTTA	CGCAAGAAAGACTTACCAGTAT	TACTGTGC		8316
Db	8372	CTTCATGAATTAAATTCAGATGTTTTA	CGCAAGAAAGACTTACCAGTAT	TACTGTGC		8431
OY	8317	TGTGCTTTTGATCTCTGCTTACCGTT	CAAAGCGGTGACAGGCCGAC	GTGCTGACCC		8376
Db	8432	TGTGCTTTTGATCTCTGCTTACCGTT	CAAAGCGGTGACAGGCCGAC	GTGCTGACCC		8491
OY	8377	CATACCTGCGAGAACAAGGGGGGGG	AATGCTGCTGACAGCCCCGCTG	TGCTCCCTC		8436
Db	8492	CATACCTGCGAGAACAAGGGGGGGG	AATGCTGCTGACAGCCCCGCTG	TGCTCCCTC		8551
OY	8437	CCCTCCCTTCTCTGGGCGAATGAAT	TGCGATTCGTGAGCCGCAATTTG	CGCAGGG		8496
Db	8552	CCCTCCCTTCTCTGGGCGAATGAAT	TGCGATTCGTGAGCCGCAATTTG	CGCAGGG		8611
OY	8497	TGTGTGTAATCTGTCAATTTACAC	CGTGTCTTATTTAAAAGCAATTA	TACTTCAAAA		8556
Db	8612	TGTGTGTAATCTGTCAATTTACAC	CGTGTCTTATTTAAAAGCAATTA	TACTTCAAAA		8671
OY	8557	AAAAAAAAAAAAAA	8571			
Db	8672	AAAAAAAAAAAAAA	8686			
<hr/>						
RESULT 9						
ADG86290						
ID	ADG86290	standard; cDNA; 8686 BP.				
XX	AC	ADG86290;				
XX	DT	11-MAR-2004 (first entry)				
DE	Human SMRT encoding cDNA SEQ ID NO:4.					
XX	KW	SMRT; silencing mediator for retinoid and thyroid hormone action;				
KW	KW	SMRT inhibitor; cytosolic; antiinflammatory; antiarthritic;				
KW	KW	antirheumatic; antisense therapy; inflammatory disorder;				
KW	KW	rheumatoid arthritis; hyperproliferative disorder; cancer; leukaemia;				
XX	OS	breast cancer; human; gene; ss.				
XX	XX	Homo sapiens.				
XX	XX					
XX	FH	Location/Qualifiers				
FT	CDS	157..7680				
FT	FT	/*tag= a				
FT	FT	/product= "SMRT"				
XX	XX	MO2003106645-A2.				
PD	24-DEC-2003.					
XX	XX					
PF	17-JUN-2003;	2003WO-US018923.				
XX	XX					
PR	17-JUN-2002;	2002US-00174014.				
XX	XX					
PA	(ISIS-) ISIS PHARM INC.					
XX	XX					
PI	Bennett CF, Freier SM, Dobie KM,					
XX	XX					
DR	WI; 2004-082184/08.					
DR	P-PSD8; ADG86291.					
DR	GENBANK; AF125672.					
XX	XX					

PT Novel antisense compound targeted to nucleic acid encoding SMRT  
PT (silencing mediator for retinoid and thyroid hormone action), useful for  
PT treating animal having disease associated with SMRT such as cancer,  
PT rheumatoid arthritis.

XX Example 13; SEQ ID NO 4; 260pp; English.

CC The present invention describes a compound (I) 8-50 nucleobases in length  
CC targeted to a nucleic acid molecule encoding SMRT (silencing mediator for  
CC retinoid and thyroid hormone action), where (I) specifically hybridizes  
CC with the nucleic acid molecule encoding SMRT and inhibits expression of  
CC SMRT. (I) specifically hybridizes with at least 8-nucleobase portion of a  
CC preferred target region on nucleic acid molecule encoding SMRT. Also  
CC described is a composition (II) comprising (I) and a carrier or diluent.  
CC (I) and (II) have cytostatic, antiinflammatory, antiarthritic and  
CC antirheumatic activities, and can be used in antisense therapy, and as  
CC SMRT expression inhibitors. (I) is useful for inhibiting the expression  
CC of SMRT in cells or tissues. (I) is also useful for treating an animal  
CC having a disease or condition associated with SMRT, e.g., inflammatory  
CC disorder such as rheumatoid arthritis; or a hyperproliferative disorder  
CC such as cancer chosen from leukemia and breast cancer, by inhibiting the  
CC expression of SMRT. (I) is useful for diagnostics, therapeutics,  
CC prophylaxis and as research reagents and kits. The present sequence  
CC encodes human SMRT, which is used in an example from the present  
CC invention.

SO Sequence 8686 BP; 1881 A; 3069 C; 2562 G; 1174 T; 0 U; 0 Other;

Query Match 97.2%; Score 8324.6; DB 12; Length 8686;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 8472; Conservative 1; Mismatches 34; Indels 88; Gaps 4;

QY 1 CANTGCTGGCTCCACACAGTTGTGGACAGAGTGGAGGCGACAGCCCGCTAACCC 60  
DB 156 CATTGCTGGCTCCACACAGCTGTGGACAGAGTGGAGGCGACAGCTGAGCCCGCTAACCC 215  
QY 61 GCCCCACAGCCTTCTTACCCAGTGCAGATCGCCCGGACGACACAGGACGTGGGCTCTCT 120  
DB 216 GCCCCACAGCCTTCTTACCCAGTGCAGATCGCCCGGACGACACAGGACGTGGGCTCTCT 215  
QY 121 GGAATGACAGACCACTCCCGGACTATGCTCCCACTGTGCGCGGCTCCATCATCCA 180  
DB 276 GGAATGACAGACCACTCCCGGACTATGCTCCCACTGTGCGCGGCTCCATCATCCA 335  
QY 181 GCCCCAGCGGCGGAGGCGCTCCCTGCTGTGATGTTCCAGCCCGGAAATGAAGGTTCCA 240  
DB 336 GCCCCAGCGGCGGAGGCGCTCCCTGCTGTGATGTTCCAGCCCGGAAATGAAGGTTCCA 395  
QY 241 GGAAGTCCACTGCGGCGAGAGTCCCACTCATACCTGCGGAGTGGGAAATGACAGAT 300  
DB 396 GGAAGTCCACTGCGGCGAGAGTCCCACTCATACCTGCGGAGTGGGAAATGACAGAT 455  
QY 301 GGAATTCATTGAAGAGACGCGCTCGGCTAGAGCTGCTGACCCCTGCTGCGACC 360  
DB 456 GGAATTCATTGAAGAGACGCGCTCGGCTAGAGCTGCTGACCCCTGCTGCGACC 515  
QY 361 GTCAACCCCTGCTGCGGCGAGGCGCTGCGGAGTGTGAAGACTCAACAGGACCGTGA 420  
DB 516 GTCAACCCCTGCTGCGGCGAGGCGCTGCGGAGTGTGAAGACTCAACAGGACCGTGA 575  
QY 421 CTTGACGGGCAAGCTGGAACCGGTGTCTCCCGCCAGCCCGCCGACACTGACCTTAGCT 480  
DB 576 CTTGACGGGCAAGCTGGAACCGGTGTCTCCCGCCAGCCCGCCGACACTGACCTTAGCT 635  
QY 481 GGAAGTGTGCGGCGACGCGTGTCCAAAGAGAGACTGATCCAGAACATGACCGCGTGA 540  
DB 636 GGAAGTGTGCGGCGACGCGTGTCCAAAGAGAGACTGATCCAGAACATGACCGCGTGA 695  
QY 541 CCGAGAGATCAACATGTGAGAGAGAGATCTTAAAGCTGAAGAAAGAGAGAGAGCT 600  
DB 696 CCGAGAGATCAACATGTGAGAGAGAGATCTTAAAGCTGAAGAAAGAGAGAGAGCT 755  
QY 601 GGAAGAGAGAGCTGCAAGCGCGGCGGCTGAGAAAGCCGCTGACCGCGCCCATCGA 660

DB 756 GGAAGAGAGAGCTGCCAAGCGCGCCAGCCTGAGAAAGCCGCTGACCCGCGCCATCGA 815  
QY 661 GTGGAAGACCCGAGCGCTGGTGCAGATCATCTACGACGAGAACCGGAAAGAGCTGAAGC 720  
DB 816 GTGGAAGACCCGAGCGCTGGTGCAGATCATCTACGACGAGAACCGGAAAGAGCTGAAGC 875  
QY 721 TGCACATCCGATTTCTGAAAGGCTTGGGCGCCCAAGTGAAGCTCCGCTGTACACACGCC 780  
DB 876 TGCACATCCGATTTCTGAAAGGCTTGGGCGCCCAAGTGAAGCTCCGCTGTACACACGCC 935  
QY 781 CTCGACACCCGCGAGTATCATGAGAACATCAAAATTAACACGAGCTGCGAAAGACT 840  
DB 936 CTCGACACCCGCGAGTATCATGAGAACATCAAAATTAACACGAGCTGCGAAAGACT 995  
QY 841 AATCTTGTACTTCAAGAGAGAGAAATCAAGCTCGGAAACAATGAGACAGAGATTCTGCA 900  
DB 996 AATCTTGTACTTCAAGAGAGAGAAATCAAGCTCGGAAACAATGAGACAGAGATTCTGCA 1055  
QY 901 GCGCTATGACACAGCTCATGGAAGGCTTGGAAAAAAGTGAAGCGCATCGAAAAACACC 960  
DB 1056 GCGCTATGACACAGCTCATGGAAGGCTTGGAAAAAAGTGAAGCGCATCGAAAAACACC 1115  
QY 961 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGGAGTACTACGAAAAAGCAAGTTCCCTGAGAT 1020  
DB 1116 CCGCGCGCGGCGCAAGAGAGCAAGGTTCGCGAGTACTACGAAAAAGCAAGTTCCCTGAGAT 1175  
QY 1021 CCGCAAGCAGCGGAGCTGACAGAGCGCATGACAGACAGGATGGGCCAGCGGCGGAGTGG 1080  
DB 1176 CCGCAAGCAGCGGAGCTGACAGAGCGCATGACAG--AGGATGGCGAGCGGCGGAGTGG 1232  
QY 1081 GCTGTCATGTGCGGCGCGCGCAGAGACAGAGTGTGAGATGATCATGAGGCTCTCTC 1140  
DB 1233 GCTGTCATGTGCGGCGCGCGCAGAGACAGAGTGTGAGATGATCATGAGGCTCTCTC 1292  
QY 1141 AGAGCAGAGAAACCTTGAGAGAGAGAGATGCGCAGCTGCGCGTATCCCGCCATGCTGTA 1200  
DB 1293 AGAGCAGAGAAACCTTGAGAGAGAGAGATGCGCAGCTGCGCGTATCCCGCCATGCTGTA 1352  
QY 1201 CGACGCTGACACAGACGCGCATCAAGTTTCATCAACATGAAACGCGCTTATGCGCCAGCCCAT 1260  
DB 1353 CGACGCTGACACAGACGCGCATCAAGTTTCATCAACATGAAACGCGCTTATGCGCCAGCCCAT 1412  
QY 1261 GAAAGTGTCAAAAGACCGCGAGGTCATGAACATGTGAAGTGAAGAGAGAGACTT 1320  
DB 1413 GAAAGTGTCAAAAGACCGCGAGGTCATGAACATGTGAAGTGAAGAGAGAGACTT 1472  
QY 1321 CCGGAGAAAGTTCAATGACAGATCCCAAGAACTTTGGCTGATGCAATTCCTGAGAG 1380  
DB 1473 CCGGAGAAAGTTCAATGACAGATCCCAAGAACTTTGGCTGATGCAATTCCTGAGAG 1532  
QY 1381 GAAAGACATGTGCTGAGTGGCTCTTATTACTTACTGACTAAGAAATGAAGAACTTAA 1440  
DB 1533 GAAAGACATGTGCTGAGTGGCTCTTATTACTTACTGACTAAGAAATGAAGAACTTAA 1592  
QY 1441 GAGCTGTGAGAGCGGAGCTATGCGCGCGGCAAGAGCGACAAACAGAGACA 1500  
DB 1593 GAGCTGTGAGAGCGGAGCTATGCGCGCGGCAAGAGCGACAAACAGAGACA 1552  
QY 1501 GCAGCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
DB 1653 GCAGCAGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1712  
QY 1561 AGATGAG 1620  
DB 1713 AGATGAG 1772  
QY 1621 CGACAAAGAGAGCCTTCTCAAGGAGAGACAGACGACCTCAGGAGAGAGACAAACAGCA 1680  
DB 1773 CGACAAAGAGAGCCTTCTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1832  
QY 1681 GAAAGAGGCTGTGCTTCCAAAGCGCGCAAACTGCAACAGCAGGAGAGAGAGAGAGAG 1740



Db	183	GAAGAGAGCTGTGGCCCTCCAAAGGCCGGAATACTGCCAACAAGCCAGGGAAGACGCAAAAG	1892
Qy	1741	CCGCATCACCCCCTCAATAGGCTTAATAGAGCCCAACAGCGAGGAGCCATCACCCCCAGCA	1800
Db	1893	CCGCATCACCCCCTCAATAGGCTTAATAGAGCCCAACAGCGAGGAGCCATCACCCCCAGCA	1952
Qy	1801	GAGGCGCGAGCTGGCCCTCCATGAGAGCTGAAATGAGAGTTCTGGCTGAGCAAGAAAGAAAT	1866
Db	1953	GAGGCGCGAGCTGGCCCTCCATGAGAGCTGAAATGAGAGTTCTGGCTGAGCAAGAAAGAAAT	2012
Qy	1861	GGAACACAGCCCAAGAAAGGTCTCTGAGAAACAGGCCGCAACTGTCGCGCATGCGCCGAT	1920
Db	2013	GGAACACAGCCCAAGAAAGGTCTCTGAGAAACAGGCCGCAACTGTCGCGCATGCGCCGAT	2072
Qy	1921	GGTGGGCTCCAGACTGTGTGCGCAGTGTAGAACTTCTACTTCAACTCAAGAAAGGCA	1980
Db	2073	GGTGGGCTCCAGACTGTGTGCGCAGTGTAGAACTTCTACTTCAACTCAAGAAAGGCA	2132
Qy	1981	GAACCTCGATAGATCTTTGACGACGACAAACTGAAATGAGAAAGGAGAAAGCGCG	2040
Db	2133	GAACCTCGATAGATCTTTGACGACGACAAACTGAAATGAGAAAGGAGAAAGCGCG	2192
Qy	2041	GAGAGAAGAAAGAAAGGCGCGCGCGCGCCAGCGAGGAGGCTCATTTCCGCGCTGTGT	2100
Db	2193	GAGAGAAGAAAGAAAGGCGCGCGCGCGCCAGCGAGGAGGCTCATTTCCGCGCTGTGT	2252
Qy	2101	GAGAGATGAGAGATGAGAGGCGTGGGCGTGAAGCGGAAATAGAGAGAGATGTTGAGAG	2160
Db	2253	GAGAGATGAGAGATGAGAGGCGTGGGCGTGAAGCGGAAATAGAGAGAGATGTTGAGAG	2312
Qy	2161	GGCTGAAGCCTTAATGCTCTTGGGAATGAGGTGCCAAGAGGGAATGCAGTGCCACGC	2220
Db	2313	GGCTGA-----AGC	2321
Qy	2221	CACGTGCAACAAAGCTCAGACACCGAGAGATCCCTCTGCTCAACTGAGGCGCGCA	2280
Db	2322	CACGTGCAACAAAGCTCAGACACCGAGAGATCCCTCTGCTCAACTGAGGCGCGCA	2381
Qy	2281	GGACACAGGGGAGATGGGCCCAAGCCCCCAGCCACCTTGAGCGCGCAAGGGCCACCCCC	2340
Db	2382	GGACACAGGGGAGATGGGCCCAAGCCCCCAGCCACCTTGAGCGCGCAAGGGCCACCCCC	2441
Qy	2341	AGGCCCAACCCCACTCAACGAGAGCATCCCGGCCCTTGAAGCCACCCGGGCTC	2400
Db	2442	AGGCCCAACCCCACTCAACGAGAGCATCCCGGCCCTTGAAGCTCAACCCGGGCTC	2501
Qy	2401	TGAAGCCACCGGAAGCCCTACGCCCCCAACGACACCCCACTGGCCCTTGACCTCCCTC	2466
Db	2502	TGAAGCCACCTTGAAGCCCTTACGCCCCCAACGACACCCCACTTCTTCAACTCTCTCC	2561
Qy	2461	TGTGTGTCCTCCACGAGAGAGAGAGAGAGAGACGCAACGCGCCCCAGTGAAGAGAGG	2520
Db	2562	TGTGTGTCCTCCACGAGAGAGAGAGAGAGAGACGCAACGCGCCCCAGTGAAGAGAGG	2621
Qy	2521	GGAGAGACAGAGCCCCCCCCGCGCTGAGAGCTGGCACTGACACAGGAAAGCCGAGGA	2580
Db	2622	GGAGAGACAGAGCCCCCCCCGCGCGCTGAAGACTGGCAGTGAACAAGGAAAGGCCGAGGA	2681
Qy	2581	GGCCGTCAGAGCGAGTGCACGAGAGAGAGCCGAGAGAGGGGCGCGCAAGGGCCAGAGACG	2640
Db	2682	GGCCGTCAGAGCGAGTGCACGAGAGAGAGCCGAGAGAGGGGCGCGCAAGGGCCAGAGACG	2741
Qy	2641	GGAAGCGCTTGAAGGACAGCGCCGAGAGGGCGCTCAAGCCAGAGAAAGAGAGGCGGAG	2700
Db	2742	GGAAGCGCTTGAAGGACAGCGCCGAGAGGGCGCTCAAGCCAGAGAAAGAGAGGCGGAG	2801
Qy	2701	CGGCAAGGACCACTGCAAGAGCTGAGGCGCCCCCGGAGCAAGCGACTTCAAGTCTAC	2760
Db	2802	CGGCAAGGACCACTGCAAGAGCTTGGGCGCCCCCGGAGCAAGCGACTTCAAGTCTAC	2861
Qy	2761	CTGCAAGTGCACACGAGTGTGATGAGAGCCGAGAGCGGCGCAACAAGACCGGCTGTCTCCC	2820
Db	2862	CTGCAAGTGCACACGAGTGTGATGAGAGCGGCGCGCGCAACAAGACCGGCTGTCTCCC	2921

QY	2821	AAAGCCCAAGCCTCTTCAACCCCGACACTGGCCACACCCCGGGCCAAATGCTCAACCCCAAGAAACC	2880
Dp	2922	AAAGCCCAAGCCTCTTCAACCCCGACTGGCCACACCCCGGGCCAAATGCTCAACCCCAAGAAACC	2981
QY	2881	ACTGGAAGCCTTAAGCAGCTGAAGCAGCAGCGAGCTGCATCCGCCCATTCAGTCAAGTCAACAA	2940
Dp	2982	ACTGGAAGCCTTAAGCAGCTGAAGCAGCAGCGAGCTGCATCCGCCCATTCAGTCAAGTCAACAA	3041
QY	2941	AGTCCATGAGCCCCCGGGGAGACGACGCTCCCAACAGCAGCTCCCTCAGCCCCCACC	3000
Dp	3042	AGTCCATGAGCCCCCGGGGAGACGACGCTCCCAACAGCAGCTCCCTCAGCCCCCACC	3101
QY	3001	GCCACCGCAAAACCTTGAGCCGAGAGCAGACGACCCCTCAGAGCCTGAGCAGCAGCACC	3060
Dp	3102	GCCACCGCAAAACCTTGAGCCGAGAGCAGACGACCCCTCAGAGCCTGAGCAGCAGCACC	3161
QY	3061	GAGGCAAGAGCAGAGAGCCCGGCAACCCCGCTCGACMAAGAG-----	3100
Dp	3162	GAGGCAAGAGCAGAGAGCCCGGCAACCCCGCTCGACMAAGAGAGAGCCTGTGTTCTT	3221
QY	3101	-----GCCTTGAGAGCCGAGGCCCAAGAGCTGCGGGGAGACCCCTTGCTTGACTTCCG	3156
Dp	3222	CCAGCCTTGAGAGCCGAGGCCCAAGAGCTGCGGGGAGACCCCTTGCTTGACTTCCG	3281
QY	3157	CTTGCCCTTCCCGCTGCGCCCCCGCTGAGGTATCAAGGCTCTCCCGCAGATGCCCCGAGACC	3216
Dp	3282	CTTGCCCTTCCCGCTGCGCCCCCGCTGAGGTATCAAGGCTCTCCCGCAGATGCCCCGAGACC	3341
QY	3217	CTCAGCCTTCTCTTCAAGCTCAACTGGTCAACCACTGCGCTCCCTGGGCTTCCATGACATGCG	3276
Dp	3342	CTCAGCCTTCTCTTCAAGCTCAACTGGTCAACCACTGCGCTCCCTGGGCTTCCATGACATGCG	3401
QY	3277	CCGACCCGCTCTGCGCGAGCCCAACCACTCATCTCAACCCGCTCCCTCATCTCTCTGCG	3336
Dp	3402	CCGACCCGCTCTGCGCGAGCCCAACCACTCATCTCAACCCGCTCCCTCATCTCTCTGCG	3461
QY	3337	CAAGCACCACCGAGCTCTTGAAGAGCAATGAGTGCATCTCCCAAGAAATGTGCGTCCA	3396
Dp	3462	CAAGCACCACCGAGCTCTTGAAGAGCAATGAGTGCATCTCCCAAGAAATGTGCGTCCA	3521
QY	3397	GCTCCAGTCCCGTACTCAAGCATGCAAGAGCCCGGTGGGCCCTGTCAACATGAGGAGCT	3456
Dp	3522	GCTCCAGTCCCGTACTCAAGCATGCAAGAGCCCGGTGGGCCCTGTCAACATGAGGAGCT	3581
QY	3457	GCCCTGCTCCATGAGACCCCAAAAAGCTTGAGCACCTTTCAGCGAGTGAAGCAGAGCAGCT	3516
Dp	3582	GCCCTGCTCCATGAGACCCCAAAAAGCTTGAGCACCTTTCAGCGAGTGAAGCAGAGCAGCT	3641
QY	3517	GTCCTCCACGGGGCCAGAGCTGGGCCACCGAGAGGCTTGGGGGTGCCACAGGCCAGAGGC	3576
Dp	3642	GTCCTCCACGGGGCCAGAGCTGGGCCACCGAGAGGCTTGGGGGTGCCACAGGCCAGAGGC	3701
QY	3577	GTCCGCTCTGAGAGAGGACAGCTCTGGGAGCTCAGTTCCGGGCGGAAGCATCAACAAAGGCAT	3636
Dp	3702	GTCCGCTCTGAGAGAGGACAGCTCTGGGAGCTCAGTTCCGGGCGGAAGCATCAACAAAGGCAT	3761
QY	3637	TCCCAAGACACAGGAGTCCCTTCGAGCAGCGCCCATCAATACCGGGCTTCATACCTCACCG	3696
Dp	3762	TCCCAAGACACAGGAGTCCCTTCGAGCAGCGCCCATCAATACCGGGCTTCATACCTCACCG	3821
QY	3697	CAGCCCAAGCTGAGCTCTGTATCAAAAGGACCATCAACAGAGTATATGGCGGAGCAGAGCC	3756
Dp	3822	CAGCCCAAGCTGAGCTCTGTATCAAAAGGACCATCAACAGAGTATATGGCGGAGCAGAGCC	3881
QY	3757	GAGTCCGCTTGAACCGAGGCGGGAGAGCAGGCTGCCAAGGCGCAGCTCATCTACGAGG	3816
Dp	3882	GAGTCCGCTTGAACCGAGGCGGGAGAGCAGGCTGCCAAGGCGCAGCTCATCTACGAGG	3941
QY	3817	CAAGAAAGGCGCAGCTTTTGTCTATGAGGGTGGCATGTCTGTGACCCAGTGTCTCAAGGA	3876
Dp	3942	CAAGAAAGGCGCAGCTTTTGTCTATGAGGGTGGCATGTCTGTGACCCAGTGTCTCAAGGA	4001



QY 3877 GAGCGCAGAGACGCTCAGACACCCCCCATGTAGACGGCGCCCCCAGACGCACTTATGA 3936  
| | | | |  
Db 4002 GGAACGGCAAGACAGCTCAGAACCCCCCATGTAGAGGGCGCCCCAAGGCACTTATGA 4061  
| | | | |  
QY 3937 CATGATGAGGGCCGCGTGGGCAAGACCATTCTTCAGCCAGCATGAAAGTCTCATGGG 3996  
| | | | |  
Db 4062 CATGATGAGGGCCGCGTGGGCAAGACCATTCTTCAGCCAGCATGAAAGTCTCATGGG 4121  
| | | | |  
QY 3997 CGGTGCATATCCCGCCGAGACGACAGCCCCCAACCTCAAGAGACACACACATCCG 4056  
| | | | |  
Db 4122 CGGTGCATATCCCGCCGAGACGACAGCCCCCAACCTCAAGAGACACACACATCCG 4181  
| | | | |  
QY 4057 CGGGTTCATACACAAAGGATCCCTCGTCTTACGTGAGAGCACAAGAGACTTACCTGGC 4116  
| | | | |  
Db 4182 CGGGTTCATACACAAAGGATCCCTCGTCTTACGTGAGAGCACAAGAGACTTACCTGGC 4241  
| | | | |  
QY 4117 TCGGAGAGCCAAAGCTCTTAAAGCGGAGGAGCAGCTCTCGCCCCCAACGCGCTTACGGGA 4176  
| | | | |  
Db 4242 TCGGAGAGCCAAAGCTCTTAAAGCGGAGGAGCAGCTCTCGCCCCCAACGCGCTTACGGGA 4301  
| | | | |  
QY 4177 CTTGACCGAGGCTTACAAAGACGAGGCTTGGGCCCCCTGAAAGTTGAACCGGCCCCATGA 4236  
| | | | |  
Db 4302 CTTGACCGAGGCTTACAAAGACGAGGCTTGGGCCCCCTGAAAGTTGAACCGGCCCCATGA 4361  
| | | | |  
QY 4237 GGGCCTGTGGCCACGCTGAAAGAGGGCGGCGCTCATCTCATGATGCCGCGCAGGA 4296  
| | | | |  
Db 4362 GGGCCTGTGGCCACGCTGAAAGAGGGCGGCGCTCATCTCATGATGCCGCGCAGGA 4421  
| | | | |  
QY 4297 GCTGGCGGACACGCGCCGAGCTGGCCCTGGCCCGCGCGCTTCAAGAGAGGCTCATCA 4356  
| | | | |  
Db 4422 GCTGGCGGACACGCGCCGAGCTGGCCCTGGCCCGCGCGCTTCAAGAGAGGCTCATCA 4481  
| | | | |  
QY 4357 GCAAGGACACCCCGCTTAAGTACACACCGCGCGCTTCACTGAGTCCAAAAGACAGA 4416  
| | | | |  
Db 4482 GCAAGGACACCCCGCTTAAGTACACACCGCGCGCTTCACTGAGTCCAAAAGACAGA 4541  
| | | | |  
QY 4417 CGTACGCTCCCTCATAGGAGACCCCGCGCGAGCTTCCACCGCTTACACCGGCTGATGT 4476  
| | | | |  
Db 4542 CGTACGCTCCCTCATAGGAGACCCCGCGCGAGCTTCCACCGCTTACACCGGCTGATGT 4601  
| | | | |  
QY 4477 GATGGCCGACGCGCCGCGCACTGAAAGTGCCTGTACAGAGAGGCTTGAAGAGCGGCG 4536  
| | | | |  
Db 4602 GATGGCCGACGCGCCGCGCACTGAAAGTGCCTGTACAGAGAGGCTTGAAGAGCGGCG 4661  
| | | | |  
QY 4537 AAGGACCCGCGACAGCTCGGGGGGCTTCAATTGCGCGCGCGCTCGCTCATTTGTGCTGA 4596  
| | | | |  
Db 4662 AAGGACCCGCGACAGCTCGGGGGGCTTCAATTGCGCGCGCGCTCGCTCATTTGTGCTGA 4721  
| | | | |  
QY 4597 GCTGGGTAGCGCGCGCAGAGCCCGCTTGAAGTGAAGACACAGGGGCAACCTTTGCGGG 4656  
| | | | |  
Db 4722 GCTGGGTAGCGCGCGCAGAGCCCGCTTGAAGTGAAGACACAGGGGCAACCTTTGCGGG 4781  
| | | | |  
QY 4657 CCACTCCACAGAGTTTCGCCGCTGACATAGCGGAGACCAAGCGCGCTGACAGAGAG 4716  
| | | | |  
Db 4782 CCACTCCACAGAGTTTCGCCGCTGACATAGCGGAGACCAAGCGCGCTGACAGAGAG 4841  
| | | | |  
QY 4717 CAGCCTTTGCTCAAGAGCATCCAGAGCGGAAAGTGAAGTGAAGTGAAGT 4776  
| | | | |  
Db 4842 CAGCCTTTGCTCAAGAGCATCCAGAGCGGAAAGTGAAGTGAAGTGAAGT 4901  
| | | | |  
QY 4777 GCGCAAGTCCCGGACAGACAGTGCAGAGACCAAGCAACCCCATCGCGCTTATGA 4836  
| | | | |  
Db 4902 GCGCAAGTCCCGGACAGACAGTGCAGAGACCAAGCAACCCCATCGCGCTTATGA 4961  
| | | | |  
QY 4837 GCACTGTCTTGGGGGCTGAGTGCCTGACCTTGAATGCAAGCAATCCCTGTGCTT 4896  
| | | | |  
Db 4962 GCACTGTCTTGGGGGCTGAGTGCCTGACCTTGAATGCAAGCAATCCCTGTGCTT 5021  
| | | | |  
QY 4897 GCACTCCACTTCATATCCCGCGGACATCCCTGTGAGCGAGCGCGTGTCTTACTTGC 4956  
| | | | |  
Db 5022 CCACTCCACTTCATATCCCGCGGACATCCCTGTGAGCGAGCGCGTGTCTTACTTGC 5081  
| | | | |  
QY 4957 CGAGACCTGTGGCCCCCAACCCACCTAACCGGACCTGTACCCACCTTACTATCCGCG 5016  
| | | | |

Db 5082 CGAGACCTGTGGCCCCCAACCCACCTAACCGGACCTGTACCCACCTTACTATCCGCG 5141  
| | | | |  
QY 5017 CTATCCCGACAGGCGCGCGCTGTAGAACCGCGAGACCATATTAAGTATCATCATCTC 5076  
| | | | |  
Db 5142 CTATCCCGACAGGCGCGCGCTGTAGAACCGCGAGACCATATTAAGTATCATCATCTC 5201  
| | | | |  
QY 5077 GGAGCAGATGACCAACAACCGGCGCAGCGCATAGGCGCGAGAGTGTATGTGAGGG 5136  
| | | | |  
Db 5202 GGAGCAGATGACCAACAACCGGCGCAGCGCATAGGCGCGAGAGTGTATGTGAGGG 5261  
| | | | |  
QY 5137 CCTTTCGCCCGAGTCTGTGCTGCACTCAATAGTGCAGGCTCCCGAGCATCAT 5196  
| | | | |  
Db 5262 CCTTTCGCCCGAGTCTGTGCTGCACTCAATAGTGCAGGCTCCCGAGCATCAT 5321  
| | | | |  
QY 5197 CGACTGTCCCAAGTCCACACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5256  
| | | | |  
Db 5322 CGACTGTCCCAAGTCCACACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5381  
| | | | |  
QY 5257 CACCGCATGAGCCGCTTGCCTTACCTCCGACCGGCGCCGAGCCCTTACGAGCCGCA 5316  
| | | | |  
Db 5382 CACCGCATGAGCCGCTTGCCTTACCTCCGACCGGCGCCGAGCCCTTACGAGCCGCA 5441  
| | | | |  
QY 5317 CAGCAGCTCCCACTTCTCCGAGAGGTCACAACAATTGACAAAACAACCAACGTC 5376  
| | | | |  
Db 5442 CAGCAGCTCCCACTTCTCCGAGAGGTCACAACAATTGACAAAACAACCAACGTC 5501  
| | | | |  
QY 5377 CTGTCTCGAGCGGAGACGAGACCGGAGTGAAGCGGGACCGGAGTGGGAGCGGGAAAA 5436  
| | | | |  
Db 5502 CTGTCTCGAGCGGAGACGAGACCGGAGTGAAGCGGGACCGGAGTGGGAGCGGGAAAA 5561  
| | | | |  
QY 5437 GTTCATCTCTCAGTCCACACAGACGCTGAGACGCAACCATTTGAGAGACTGTATCAGA 5496  
| | | | |  
Db 5562 GTTCATCTCTCAGTCCACACAGACGCTGAGACGCAACCATTTGAGAGACTGTATCAGA 5621  
| | | | |  
QY 5497 GCAAGACAGCGGCAAGCAGCGGAGCAGCGCGGAGTGGGGGAGCAGCAGCGCCCGCG 5556  
| | | | |  
Db 5622 GCAAGACAGCGGCAAGCAGCGGAGCAGCGCGGAGTGGGGGAGCAGCAGCGCCCGCG 5681  
| | | | |  
QY 5557 CTTCACCTCCATGCCCCACACAGACTGCGCCATCTTCCCTCGGACCCAGAGTGCCTTCA 5616  
| | | | |  
Db 5682 CTTCACCTCCATGCCCCACACAGACTGCGCCATCTTCCCTCGGACCCAGAGTGCCTTCA 5741  
| | | | |  
QY 5617 GCAAGACCCCAATGAGCTTCAACAACAGGATGAAGGGTATCATACCGCTGTGAGGC 5676  
| | | | |  
Db 5742 GCAAGACCCCAATGAGCTTCAACAACAGGATGAAGGGTATCATACCGCTGTGAGGC 5801  
| | | | |  
QY 5677 CAGCAAGCCCAAGTCTGTAGGTTCACACTTCACTTCAACCGTTCGCGCAGCTGCAC 5736  
| | | | |  
Db 5802 CAGCAAGCCCAAGTCTGTAGGTTCACACTTCACTTCAACCGTTCGCGCAGCTGCAC 5861  
| | | | |  
QY 5737 ATTCCCACTTGCACACCACTGCCCCACTGGGGCGGACCTTCGATGGGGTCTTACCTT 5796  
| | | | |  
Db 5862 ATTCCCACTTGCACACCACTGCCCCACTGGGGCGGACCTTCGATGGGGTCTTACCTT 5921  
| | | | |  
QY 5797 CATGAGACCCGCTTGTGCGCCCAAGAGGCCCCCGGGGTGCGCCGCGCAGAGGCGCGCG 5856  
| | | | |  
Db 5922 CATGAGACCCGCTTGTGCGCCCAAGAGGCCCCCGGGGTGCGCCGCGCAGAGGCGCGCG 5981  
| | | | |  
QY 5857 AGCAGACACCGGCAATGCTTCTGCGCAAGCCCGCAGCGCGCTCGGCGTGAAGCCCG 5916  
| | | | |  
Db 5982 AGCAGACACCGGCAATGCTTCTGCGCAAGCCCGCAGCGCGCTCGGCGTGAAGCCCG 6041  
| | | | |  
QY 5917 CTCTTCCCCAGCAAGGCTCGAGGCCCCCGGCCCTTATGTCCTTCTGTCTTGTGCCACGC 5976  
| | | | |  
Db 6042 CTCTTCCCCAGCAAGGCTCGAGGCCCCCGGCCCTTATGTCCTTCTGTCTTGTGCCACGC 6101  
| | | | |  
QY 5977 CACATTCGCGCGCAACCCCTGCGAAGAACTGCACTTACACAGCAGCGCGGACCGGCG 6036  
| | | | |  
Db 6102 CACATTCGCGCGCAACCCCTGCGAAGAACTGCACTTACACAGCAGCGCGGACCGGCG 6161  
| | | | |  
QY 6037 GCGGCACTGTGCTGCGCTCGGACCGGACCGGAAAAAGCTCAAGTAAACCTTTTTC 6096  
| | | | |

Db 6162 GAGCCACCTGCTGCGCTGCGACCAGCCGCAACCGGAAAAAGACTCAAGTAACCTTTTC 6221  
QY 6097 CATCCAGGAATCGAAATCCGTTCTCTGGGTTACACGCGACAGCTACAGCCCGGAAG 6156  
Db 6222 CATCCAGGAATCGAAATCCGTTCTCTGGGTTACACGCGACAGCTACAGCCCGGAAG 6281  
QY 6157 GGTGAGCCCGTCAAGCCTGTGTAGCTCACCCAGTCTGACCCAGACAGAGGGCTCCCAA 6216  
Db 6282 GGTGAGCCCGTCAAGCCTGTGTAGCTCACCCAGTCTGACCCAGACAGAGGGCTCCCAA 6341  
QY 6217 GCACCTGGAAGAGCTCGAACAGGCACTTGAGGGGGAGCTGCGGCCCAAGCAGCGAG 6276  
Db 6342 GCACCTGGAAGAGCTCGAACAGGCACTTGAGGGGGAGCTGCGGCCCAAGCAGCGAG 6401  
QY 6277 CCCCGTGAAGCTTGGCGGGAGGCGCCCACTCCCAACCTTGCGCGCTGTGAGAG 6336  
Db 6402 CCCCGTGAAGCTTGGCGGGAGGCGCCCACTCCCAACCTTGCGCGCTGTGAGAG 6461  
QY 6337 CCAAGCCTGTCAAGCCGCTGTCTCCAGACCGCCCAAGGGGTCAAAAGTCAACAGCGGT 6396  
Db 6462 CCAAGCCTGTCTCAAGCCGCTGTCTCCAGACCGCCCAAGGGGTCAAAAGTCAACAGCGGT 6521  
QY 6397 GGTCAACCCGCGCCAGCAATCAGTGAAGGTCTATCAACAGGACTACACCCGCAACACCC 6456  
Db 6522 GGTCAACCCGCGCCAGCAATCAGTGAAGGTCTATCAACAGGACTACACCCGCAACACCC 6581  
QY 6457 ACAGAGCTCAAGCGACCCCTGCGCCGCCCTCTACTCTTCCTCGGGGCGAGCTGCGC 6516  
Db 6582 ACAGAGCTCAAGCGACCCCTGCGCCGCCCTCTACTCTTCCTCGGGGCGAGCTGCGC 6641  
QY 6517 GGTCTGGAAGCTCCGCGCGCCCAACAGTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 6576  
Db 6642 GGTCTGGAAGCTCCGCGCGCCCAACAGTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 6701  
QY 6577 CCGCGCGCGGTGCTCCCGCAAGCGAAAGGGGCAAGAGGTCTCCAGAGCCAAAGAGAG 6636  
Db 6702 CCGCGCGCGGTGCTCCCGCAAGCGAAAGGGGCAAGAGGTCTCCAGAGCCAAAGAGAG 6761  
QY 6637 GTCCGTCTTGGGTGTGTGTGAGAGCGGTATTGAACTGTGTCCCAACCGGAGGAGTGA 6696  
Db 6762 GTCCGTCTTGGGTGTGTGTGAGAGCGGTATTGAACTGTGTCCCAACCGGAGGAGTGA 6821  
QY 6697 GAGGCAAGGGCACTCCCGAGTGTGTGTATCCCGCTGTGTACCGGAGTGGGGAAACAGAC 6756  
Db 6822 GAGGCAAGGGCACTCCCGAGTGTGTGTATCCCGCTGTGTACCGGAGTGGGGAAACAGAC 6881  
QY 6757 GAGGCGCAAGGAGTGGGTCCAAAGTCTCAAGGCAACACAGCGAGCGGCGGCTTCTT 6816  
Db 6882 GAGGCGCAAGGAGTGGGTCCAAAGTCTTCAAGGCAACACAGCGAGCGGCGGCTTCTT 6941  
QY 6817 CAGCAAGCTGACCGAGAGCAACTCCGCCATGTGTCAAGTCCAAAGAGCAAGAGATCAACA 6876  
Db 6942 CAGCAAGCTGACCGAGAGCAACTCCGCCATGTGTCAAGTCCAAAGAGCAAGAGATCAACA 7001  
QY 6877 GAAAGCTGAACCCCAACACCGGAGTGAAGCTGAATTAATATCAGCCAGCTGGAGCGAG 6936  
Db 7002 GAAAGCTGAACCCCAACACCGGAGTGAAGCTGAATTAATATCAGCCAGCTGGAGCGAG 7061  
QY 6937 GATCTTCAATATGCCCCGATCAACCGGAAACAGGCTTTATGACCTTTATGAAGCAGCGGT 6996  
Db 7062 GATCTTCAATATGCCCCGATCAACCGGAAACAGGCTTTATGACCTTTATGAAGCAGCGGT 7121  
QY 6997 GCAAGGACATGCGACAGCAACAATGAGGCTGGAAGGCAATTAATTAAGAGGCACTATGGG 7056  
Db 7122 GCAAGGACATGCGACAGCAACAATGAGGCTGGAAGGCAATTAATTAAGAGGCACTATGGG 7181  
QY 7057 TAAATATGACAGTGGGAAAGAGTCCCGCGCTCAGCGCAATGCTTTTAAACCTCTGAA 7116  
Db 7182 TAAATATGACAGTGGGAAAGAGTCCCGCGCTCAGCGCAATGCTTTTAAACCTCTGAA 7241  
QY 7117 TGCCAGTCCAGGCTGCGCGCTGTATGCCCCATACCGGTGTGACGAGAGGAGTGAACA 7176  
Db 7242 TGCCAGTCCAGGCTGCGCGCTGTATGCCCCATACCGGTGTGACGAGAGGAGTGAACA 7301

QY 7177 CACACTCACTTCGAGAGTGGCGCGGAGAGCCAAAGTCTTGTGACAGACCGAGCGG 7236  
Db 7302 CACACTCACTTCGAGAGTGGCGCGGAGAGCCAAAGTCTTGTGACAGACCGAGCGG 7361  
QY 7237 AAAAGCCAAAGTCCCGGCGCGGCGCTGCAATCTGGGGAACCGGCAACCTCTGTCTCTC 7296  
Db 7362 AAAAGCCAAAGTCCCGGCGCGGCGCTGCAATCTGGGGAACCGGCAACCTCTGTCTCTC 7421  
QY 7297 AGTGAAGTGAAGGAGACTGCAACCGCGGAGCGCGCTCAACCAACCGCTGTGGAGGA 7356  
Db 7422 AGTGAAGTGAAGGAGACTGCAACCGCGGAGCGCGCTCAACCAACCGCTGTGGAGGA 7481  
QY 7357 CAGGCGCTGTTCGAGAGTTCACGCGATTCCTTCAACCCCTGATCATGCGCTGCA 7416  
Db 7482 CAGGCGCTGTTCGAGAGTTCACGCGATTCCTTCAACCCCTGATCATGCGCTGCA 7541  
QY 7417 GCGGAGTGTCAATGCTTCCCAACCGGCGCTCCCGCGGCGAGCGGCGCTCTGCG 7476  
Db 7542 GCGGAGTGTCAATGCTTCCCAACCGGCGCTCCCGCGGCGAGCGGCGCTCTGCG 7601  
QY 7477 TGCGCCCGACCAAGCTGGAGAGAGAGCCCAAGGCACTGCTGTGCTGAGTACGAGAC 7536  
Db 7602 TGCGCCCGACCAAGCTGGAGAGAGAGCCCAAGGCACTGCTGTGCTGAGTACGAGAC 7661  
QY 7537 ACTCTCCAGACGAGTGAATCAGAACAGGCGGGGGGGGGCGGGCGGTGEBRSPBRUUG 7596  
Db 7662 ACTCTCCAGACGAGTGAATCAGAACAGGCGGGGGGGGGCGGGCGGTGEBRSPBRUUG 7711  
QY 7597 TCAAGTCCCAAGAGAGCCACAGGAAGGCGCTGTGAGAGAGCGGCGGCTGCGGACCTCCCC 7656  
Db 7712 TCAAGTCCCAAGAGAGCCACAGGAAGGCGCTGTGAGAGAGCGGCGGCTGCGGACCTCCCC 7771  
QY 7657 AACCAAGGAGAGAGCGCGTGAAGTCCGCGCTGCATTCATCTGTCTGCTCAAGGCG 7716  
Db 7772 AACCAAGGAGAGAGCGCGTGAAGTCCGCGCTGCATTCATCTGTCTGCTCAAGGCG 7831  
QY 7717 GCATCTTTCCTGTCTTAAAGCCTTAACTTAAGCTTCCGCGCGCGGCTGCTGTGAGA 7776  
Db 7832 GCATCTTTCCTGTCTTAAAGCCTTAACTTAAGCTTCCGCGCGCGGCTGCTGTGAGA 7891  
QY 7777 CCTTACTGAAGGAGATGTTTAACTGTGTGTGAGGAAGGAGGAGGAGCGGAGGAGGAGG 7836  
Db 7892 CCTTACTGAAGGAGATGTTTAACTGTGTGTGAGGAAGGAGGAGGAGGAGGAGGAGG 7951  
QY 7837 GCAAGGCAAGGCGGTGTGAGGCAACACAGGCGGCAAGGCGGCGGAGGAGGAGGAGGAGG 7896  
Db 7952 GCAAGGCAAGGCGGTGTGAGGCAACACAGGCGGCAAGGCGGCGGAGGAGGAGGAGGAGG 8011  
QY 7897 GATGACCAAGGCACTTCAAGGCACTGCTTCCCGAATGCAATTTGGAACCAAGTCTTAA 7956  
Db 8012 GATGACCAAGGCACTTCAAGGCACTGCTTCCCGAATGCAATTTGGAACCAAGTCTTAA 8071  
QY 7957 CTGAGCTGCAAGCGCGCGCGCGCTTCCCTCCGCTTCCCATCCCGCTTACGCGCTTGAAG 8016  
Db 8072 CTGAGCTGCAAGCGCGCGCGCGCTTCCCTCCGCTTCCCATCCCGCTTACGCGCTTGAAG 8131  
QY 8017 ATGAGCGAGGCGCTGTACAGCGCGCGAGTGTGCTGCTGCGGCTCCGACAGCTGCGCCA 8076  
Db 8132 ATGAGCGAGGCGCTGTACAGCGCGCGAGTGTGCTGCTGCGGCTCCGACAGCTGCGCCA 8191  
QY 8077 GCAACGAGATTTGCTGGAACCAAGTCAAGGCAAGTGGGCGGACAAAGGCGCAAGTGG 8136  
Db 8192 GCAACGAGATTTGCTGGAACCAAGTCAAGGCAAGTGGGCGGACAAAGGCGCAAGTGG 8251  
QY 8137 GCTTGGGGGGAACGAGTCTCCAGAGACTGACCTGTTTTTCAACATCGTTGCGGAG 8196  
Db 8252 GCTTGGGGGGAACGAGTCTCCAGAGACTGACCTGTTTTTCAACATCGTTGCGGAG 8311  
QY 8197 GCGTGGGGAAGGAAGGAGATGAATGATGTTGTTTCAAGGCTATATTTTGTATAC 8256  
Db 8312 GCGTGGGGAAGGAAGGAGATGAATGATGTTGTTTCAAGGCTATATTTTGTATAC 8371

QY 8257 CTTCATGATTAATTAATCAGATGTTTTACGCAAGAGAGACTTACCCAGTATTACTCTGC 8316  
| | | | |  
DB 8372 CTTCATGATTAATTAATCAGATGTTTTACGCAAGAGAGACTTACCCAGTATTACTCTGC 8431  
| | | | |  
QY 8317 TGTGCTTTGATCTCTGCTTACCGTTCAAGAGCGTGTGACGCCGACAGTGGTGAACC 8376  
| | | | |  
DB 8432 TGTGCTTTGATCTCTGCTTACCGTTCAAGAGCGTGTGACGCCGACAGTGGTGAACC 8491  
| | | | |  
QY 8377 CATCATCCGACAGACCAAGGGGGGGGAGCTGCTGTCACGCCCGCTGTCTCTCC 8436  
| | | | |  
DB 8492 CATCATCCGACAGACCAAGGGGGGGGAGCTGCTGTCACGCCCGCTGTCTCTCC 8551  
| | | | |  
QY 8437 CTTCCCTTCTTGGGAGAGATTAATCAGATGCGTATTCTGTGGCCGCCATTTGCGAGGG 8496  
| | | | |  
DB 8552 CTTCCCTTCTTGGGAGAGATTAATCAGATGCGTATTCTGTGGCCGCCATTTGCGAGGG 8611  
| | | | |  
QY 8497 TGTGCTTTGATCTCTGCTTACCGTTCAAGAGCGTGTGACGCCGACAGTGGTGAACC 8556  
| | | | |  
DB 8612 TGTGCTTTGATCTCTGCTTACCGTTCAAGAGCGTGTGACGCCGACAGTGGTGAACC 8671  
| | | | |  
QY 8557 AAAAAAAAAAAAAA 8571  
| | | | |  
DB 8672 AAAAAAAAAAAAAA 8686  
| | | | |  
RESULT 10  
ADQ18920  
ID ADQ18920 standard; DNA; 8686 BP.  
XX  
AC ADQ18920;  
XX  
DT 26-AUG-2004 (first entry)  
XX  
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 1739.  
XX  
KM soft tissue sarcoma; cytosstatic; gene therapy; vaccine; screening; human;  
XX  
OS Homo sapiens.  
XX  
PN WO2004048938-A2.  
XX  
PD 10-JUN-2004.  
XX  
PR 26-NOV-2003; 2003WO-US038193.  
XX  
PR 26-NOV-2002; 2002US-0429739P.  
XX  
PA (PROT-) PROTEIN DESIGN LABS INC.  
XX  
PI Aziz N, Ginsburg WM, Zlotnik A;  
XX  
XX WPI; 2004-441208/41.  
DR  
XX  
XX  
PT Early detection of soft tissue sarcoma comprises determining expression  
PT of a gene in a first soft tissue sample and a normal soft tissue sample  
PT and comparing the gene expression, also useful in treating soft tissue  
PT sarcoma.  
PS  
PS Example 2; SEQ ID NO 1739; 210pp; English.  
CC The invention relates to a novel method for detecting soft tissue sarcoma  
CC which comprises obtaining a first soft tissue sample from an individual  
CC and a normal soft tissue sample from the same or different individual,  
CC determining the expression of a gene in both samples and comparing the  
CC expression of the gene in both soft tissue samples, where a higher level  
CC of protein expression in the first soft tissue sample indicates the  
CC presence of soft tissue sarcoma. The method of the invention has  
CC cytostatic applications and may be useful for detecting soft tissue  
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic  
CC acid sequences may be useful in diagnostic and screening applications.  
CC The current sequence is that of a human soft tissue sarcoma-upregulated  
CC DNA of the invention. The current sequence is not shown within the

CC specification per se but was submitted in CD format by the inventor.  
XX  
SQ Sequence 8686 BP; 1881 A; 3069 C; 2562 G; 1174 T; 0 U; 0 Other;  
Query Match 97.2%; Score 8324.6; DB 12; Length 8686;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 8472; Conservative 1; Mismatches 34; Indels 88; Gaps 4;  
QY 1 CATGTGGGCTCCACACAGCTTGTGACAGAGGTGAGGGCCAGTGAAGCCGCTAACC 60  
| | | | |  
DB 156 CATGTGGGCTCCACACAGCTTGTGACAGAGGTGAGGGCCAGTGAAGCCGCTAACC 215  
| | | | |  
QY 61 GCCCCACAGCTTTCTTACCCAGTGCAGATCGCCCGACACACAGAGTGGGCTCT 120  
| | | | |  
DB 216 GCCCCACAGCTTTCTTACCCAGTGCAGATCGCCCGACACAGAGTGGGCTCT 275  
| | | | |  
QY 121 GGAATACAGACCACTCCCGGAGTATGCTTCCACTGTGCGGGCTCCATCTCA 180  
| | | | |  
DB 276 GGAATACAGACCACTCCCGGAGTATGCTTCCACTGTGCGGGCTCCATCTCA 335  
| | | | |  
QY 181 GCCCCAGGGCGGAGGCGCTCCCTGCTGTGAGTTCCAGCCCGGATGAACGGTCCCA 240  
| | | | |  
DB 336 GCCCCAGGGCGGAGGCGCTCCCTGCTGTGAGTTCCAGCCCGGATGAACGGTCCCA 395  
| | | | |  
QY 241 GGAAGTCCACTGCGGCGAGAGTCCCACTCATACCTGCCAGCTGGGAAATCAGAGAT 300  
| | | | |  
DB 396 GGAAGTCCACTGCGGCGAGAGTCCCACTCATACCTGCCAGCTGGGAAATCAGAGAT 455  
| | | | |  
QY 301 GGAATTCATTGAAGAGCAGCGCTTGGGTAAGCTGTGCTGACCCCTGTGCGAAC 360  
| | | | |  
DB 456 GGAATTCATTGAAGAGCAGCGCTTGGGTAAGCTGTGCTGACCCCTGTGCGAAC 515  
| | | | |  
QY 361 GTCACTCCCTGTGGCCACAGGGCCAGCTGCGGGAATGTGAAGCTTCAAGAGCCGTAG 420  
| | | | |  
DB 516 GTCACTCCCTGTGGCCACAGGGCCAGCTGCGGGAATGTGAAGCTTCAAGAGCCGTAG 575  
| | | | |  
QY 421 CCTGACGGGCAAGCTGGAACCGGTGTCTCCCGCAGCCCGCCGACACTGACCTTGAAGCT 480  
| | | | |  
DB 576 CCTGACGGGCAAGCTGGAACCGGTGTCTCCCGCAGCCCGCCGACACTGACCTTGAAGCT 635  
| | | | |  
QY 481 GGAAGTGTGTCGCGCCACAGGCTGTCCAGAGAGAGCTGATCCAGAACTGACCGCTGCA 540  
| | | | |  
DB 636 GGAAGTGTGTCGCGCCACAGGCTGTCCAGAGAGAGCTGATCCAGAACTGACCGCTGCA 695  
| | | | |  
QY 541 CCGAGAGATCAACATGCTAGACAGAGATCTTAAGCTGAAGAGAAACAGACAGAGCT 600  
| | | | |  
DB 696 CCGAGAGATCAACATGCTAGACAGAGATCTTAAGCTGAAGAGAAACAGACAGAGCT 755  
| | | | |  
QY 601 GGAAGAGAGAGCTGCAAGCGCCGCGAGCTGGAAGCCGCTGTCAACCGCCCATCGA 660  
| | | | |  
DB 756 GGAAGAGAGAGCTGCAAGCGCCGCGAGCTGGAAGCCGCTGTCAACCGCCCATCGA 815  
| | | | |  
QY 661 GTGGAAGACCGGACAGCTGTGTCAGATCTACGAGAGAAACCGAGAAAGGCTGAAGCT 720  
| | | | |  
DB 816 GTGGAAGACCGGACAGCTGTGTCAGATCTACGAGAGAAACCGAGAAAGGCTGAAGCT 875  
| | | | |  
QY 721 TGCACATCGGAATTTGGAAGGCTGGGGCCCAAGTGTGAGCTGCGCTTACCAACAAGCC 780  
| | | | |  
DB 876 TGCACATCGGAATTTGGAAGGCTGGGGCCCAAGTGTGAGCTGCGCTTACCAACAAGCC 935  
| | | | |  
QY 781 CTCGACACCGGCGAGATCATGAGAACATCAAAATTAACAGAGCGATCGGAAGAGCT 840  
| | | | |  
DB 936 CTCGACACCGGCGAGATCATGAGAACATCAAAATTAACAGAGCGATCGGAAGAGCT 995  
| | | | |  
QY 841 AATCTTTGATCTCAAGAGAGGAATCAAGCTGTGGAACATGGAACAGAGAGTCTTGCA 900  
| | | | |  
DB 996 AATCTTTGATCTCAAGAGAGGAATCAAGCTGTGGAACATGGAACAGAGAGTCTTGCA 1055  
| | | | |  
QY 901 GGGCTATGACCAAGCTCATATGAGGCGCTTGAATAAAGGTGAGCGGATGAATAAACC 960  
| | | | |  
DB 1056 GGGCTATGACCAAGCTCATATGAGGCGCTTGAATAAAGGTGAGCGGATGAATAAACC 1115  
| | | | |  
QY 961 GGGCGGGCGGGCCAAAGAGAGAGAGGTGCGGAGTACTAGAAAGAGTCTCGTAGAT 1020  
| | | | |

Db 1116 CCGGGGGGGCCAAAGAGCAAGGTTCCGAGTACTACGAAGCAGTTCCCTGAGAT 1175  
Qy 1021 CCGAAGCAGCGCGAGCTGCGAGAGCGATGCGAGAGCGAGGTCAGCGGCGAGTGG 1080  
Db 1176 CCGCAAGCAGCGCGAGCTGCGAGAGCGATGCGAG--AGGTTGGGCGAGCGGCGAGTGG 1232  
Qy 1081 GCTGTCAATGTGCGCGCGCGAGCGAGCGAGGTTCAAGATCATCGATGAGCTCTTC 1140  
Db 1233 GCTGTCAATGTGCGCGCGCGAGCGAGCGAGGTTCAAGATCATCGATGAGCTCTTC 1292  
Qy 1141 AGAGCAGAGAACTTGGAGAGACAGATGCGCGAGCTGCGCTGATTCGCGCCATGCTGTA 1200  
Db 1293 AGAGCAGAGAACTTGGAGAAACAGATGCGCGAGCTGCGCTGATTCGCGCCATGCTGTA 1352  
Qy 1201 CGACGCTGACGAGCGCGATCAAGTTCAATCAACATGAAAGGAGCTTATGCGCGAGCCCAT 1260  
Db 1353 CGACGCTGACGAGCGCGATCAAGTTCAATCAACATGAAAGGAGCTTATGCGCGAGCCCAT 1412  
Qy 1261 GAAGGTGTAACAAGACCGCGAGGTCAATGAACATGTGAGTGAAGAGAGAGACCTT 1320  
Db 1413 GAAGGTGTAACAAGACCGCGAGGTCAATGAACATGTGAGTGAAGAGAGAGACCTT 1472  
Qy 1331 CCGGAGAAAGTTCAATGACAGATCCCAAGAACTTTGGCTGATCGATCTTCTGGAAG 1380  
Db 1473 CCGGAGAAAGTTCAATGACAGATCCCAAGAACTTTGGCTGATCGATCTTCTGGAAG 1532  
Qy 1381 GAAGCAGTGGCTGAGTGGCTCTCTATTAATACCTGACTAAGAAAGATGAGAACTATA 1440  
Db 1533 GAAGCAGTGGCTGAGTGGCTCTCTATTAATACCTGACTAAGAAAGATGAGAACTATA 1592  
Qy 1441 GAGCTGTGAGACGGAAGCTATCGCGCGCGCGCAAGAGCCAGCAGCAACAACAGCAGA 1500  
Db 1593 GAGCTGTGAGACGGAAGCTATCGCGCGCGCGCAAGAGCCAGCAGCAACAACAGCAGA 1652  
Qy 1501 GCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
Db 1653 GCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1712  
Qy 1561 AGATGAGAAAGAGAGAAAGAGAGCGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
Db 1713 AGATGAGAAAGAGAGAAAGAGAGCGGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1772  
Qy 1621 CGACAAAGAGAGACTCTCTCAAGAGAGAGACAGACGACCTCAGGAGAGAGACAGACGAC 1680  
Db 1773 CGACAAAGAGAGACTCTCTCAAGAGAGAGACAGACGACCTCAGGAGAGAGACAGACGAC 1832  
Qy 1681 GAAGAGGCTGTGGCTCTCAAGAGCGCAAACTGCGCAACAGCAGGAGAGAGCGCAAGG 1740  
Db 1833 GAAGAGGCTGTGGCTCTCAAGAGCGCAAACTGCGCAACAGCAGGAGAGAGCGCAAGG 1892  
Qy 1741 CGGCATCAACCGCTCAATGGCTAATGAGGCAACAGCGAGAGAGGACATCACCCCGACGA 1800  
Db 1893 CGGCATCAACCGCTCAATGGCTAATGAGGCAACAGCGAGAGAGGACATCACCCCGACGA 1952  
Qy 1801 GAGCGCCGAGCTGAGCTCTCCATGAGCTGAATGAGAGTTCTCGCTGGAAGAGAGAGAAAT 1860  
Db 1953 GAGCGCCGAGCTGAGCTCTCCATGAGCTGAATGAGAGTTCTCGCTGGAAGAGAGAGAAAT 2012  
Qy 1861 GGAACACGCGAAGAAAGTCTCTGGAACAACGCGCGCAACTGTGGCCATGCGCCCGAT 1920  
Db 2013 GGAACACGCGAAGAAAGTCTCTGGAACAACGCGCGCAACTGTGGCCATGCGCCCGAT 2072  
Qy 1921 GGTGGGCTTCAAGACTGTGTGCGAGGTAAAGAACTTCACTTCAACTAAGAGAGGCA 1980  
Db 2073 GGTGGGCTTCAAGACTGTGTGCGAGGTAAAGAACTTCACTTCAACTAAGAGAGGCA 2132  
Qy 1981 GAACCTCGATGAGATCTTGACAGCAGACAAAGCTGAATGAGAGAGAGAGAGAGAGAGAG 2040  
Db 2133 GAACCTCGATGAGATCTTGACAGCAGACAAAGCTGAATGAGAGAGAGAGAGAGAGAGAG 2192  
Qy 2041 GAGGAAGAAAGAAAGAGCGCGCGCGCGCGCGCGAGAGAGGCTGCATTCGCCCGCTGGT 2100

Db 2193 GAGGAAGAAAGAAAGAGCGCGCGCGCGCGCGCGAGAGAGGCTGCATTCGCCCGCTGGT 2252  
Qy 2101 GAGAGATGAGAGATGAGAGGCTCGGCGGTGACGGAATGAGAGAGATGTTGAGAGA 2160  
Db 2253 GAGAGATGAGAGATGAGAGGCTCGGCGGTGACGGAATGAGAGAGATGTTGAGAGA 2312  
Qy 2161 GAGTGAAGCTTTACATGCGCTCTGGGAATGAGGTGCCAGAGGGAGAAATGACGTGCCAGC 2220  
Db 2313 GAGTGA-----AGC 2321  
Qy 2221 CACTGTCAACAACAGCTGACACCGAGAGATATCCCTCTCTCTCACTAGAGAGCGGCA 2280  
Db 2322 CACTGTCAACAACAGCTGACACCGAGAGATATCCCTCTCTCTCACTAGAGAGCGGCA 2381  
Qy 2281 GGAACAAGGAGAGATGAGGCGGCAAGCGCCAGCAGCAGCTTGGGCGCGAGCGGCGCAGCC 2340  
Db 2382 GGAACAAGGAGAGATGAGGCGGCAAGCGCCAGCAGCAGCTTGGGCGCGAGCGGCGCAGCC 2441  
Qy 2341 AGGCCACCCACCCCAACAAGAGACATCCCGGCGCGCATTTAGAGCCACCCGCGCTC 2400  
Db 2442 AGGCCACCCACCCCAACAAGAGACATCCCGGCGCGCATTTAGAGCTCACCCGCGCTC 2501  
Qy 2401 TGAAGCGACCGGAGCGCTTACGCGCGCGCGCGAGACATCCCATGCGCTGACACTCTCC 2460  
Db 2502 TGAAGCGACCTTACGCGCGCGCGCGAGACATCCCATTTCTTCACTCTCTCC 2561  
Qy 2461 TGTGTCTCCCAAGAGAGAGAGAGAGACGCGACAGCGCGCCAGTGAAGAGAG 2520  
Db 2562 TGTGTCTCCCAAGAGAGAGAGAGAGACGCGACAGCGCGCCAGTGAAGAGAG 2621  
Qy 2521 GAGAGAGCAAGAGCGCGCGCGCGGTGAGAGTGGCAATGACACAGGAAAGCGGAGAG 2580  
Db 2622 GAGAGAGCAAGAGCGCGCGCGCGGTGAGAGTGGCAATGAGACAGAGAAAGCGGAGAG 2681  
Qy 2581 GCGCGTCAAGAGCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
Db 2682 GCGCGTCAAGAGCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2741  
Qy 2641 GAGAGCGCTGAGAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
Db 2742 GAGAGCGCTGAGAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2801  
Qy 2701 CGGCAAGGCTCACTGCGCAAGAGCTCGGAGCGCTCCCGAGACAGGACTCTCAATGTCTAC 2760  
Db 2802 CGGCAAGGCTCACTGCGCAAGAGCTCGGAGCGCTCCCGAGACAGGACTCTCAATGTCTAC 2861  
Qy 2761 CTGCAATGCAACGAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820  
Db 2862 CTGCAATGCAACGAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2921  
Qy 2821 AAGGCCAGCTCTCTCAACCGGACTGAGCGAGCGCGCGGCGCAATGCTTCAACCCAGAGAG 2880  
Db 2922 AAGGCCAGCTCTCTCAACCGGACTGAGCGAGCGCGCGGCGCAATGCTTCAACCCAGAGAG 2981  
Qy 2881 ACTGGACTGAAAGAGCTGAAAGCAGCGAGCGCTGCAATTCCTCCCATTCAGGTACCA 2940  
Db 2982 ACTGGACTGAAAGAGCTGAAAGCAGCGAGCGCTGCAATTCCTCCCATTCAGGTACCA 3041  
Qy 2941 AGTCCATGAGCGCGCGGAG 3000  
Db 3042 AGTCCATGAGCGCGCGGAG 3101  
Qy 3001 GCAACCGCAAAACCTGACGCGGAG 3060  
Db 3102 GCAACCGCAAAACCTGACGCGGAG 3161  
Qy 3061 GGGCAAGAGCAGAGAGCGCGGAG 3100  
Db 3162 GGGCAAGAGCAGAGAGCGCGGAG 3221  
Qy 3101 -----GCTTGGCAGCGGAG 3156  
Db 3222 CCAGCTTTCGAGCGGAG 3281

QY	3157	CTGCGCCCTTCCCCCGTGGCCCCCCCCCGTGAAGTGTCAAGGCTTCGCCGATGCCGCCGGAACCC	3216
Dp	3282	CCTGCCCCCTTCCCCCGGGCCCCCCCCCGGAGAGTGTCAAGGCTTCGCCGATGCCGCCGGAACCC	3341
QY	3217	CTCAGGCTTCTCTCAAGGCTTCCACCTGGTCAACCACTGGCCCCCTGGGCTTCCTCAATGACATGCG	3276
Dp	3342	CTCAGGCTTCTCTCAAGGCTTCCACCTGGTCAACCACTGGCCCCCTGGGCTTCCTCAATGACATGCG	3401
QY	3277	CCGGCCCGTCCCTGGCCGGGCCCAACCCACATCTCAACCCGCTCCCTCATCTCTCTGC	3336
Dp	3402	CCGGCCCGTCCCTGGCCGGGCCCAACCCACATCTCAACCCGCTCCCTCATCTCTCTGC	3461
QY	3337	CAAGCAACCCCAAGCCGCTCTCGAAGGCAATATGTGSCATCTCCCAAGGAATGTGTCCTCA	3396
Dp	3462	CAAGCAACCCCAAGCCGCTCTCGAAGGCAATATGTGSCATCTCCCAAGGAATGTGTCCTCA	3521
QY	3397	GCTCCACGTCCTCGTACTCGAGACATGCGCAAGGCCCCCGGTGGGCCCTGTCACTATGGGACT	3456
Dp	3522	GCTCCACGTCCTCGTACTCGAGACATGCGCAAGGCCCCCGGTGGGCCCTGTCACTATGGGACT	3581
QY	3457	GCCCCCTGGCCATGGAACCCCAAAAGCTGGCACCTTCAGCCGAGTGAAGCAAGAGCACT	3516
Dp	3582	GCCCCCTGGCCATGGAACCCCAAAAGCTGGCACCTTCAGCCGAGTGAAGCAAGAGCACT	3641
QY	3517	GTCCCCACGGGGGCAAGGCTGGGGCCACCGGAAGGCTGGGGGGTGCCCAAGGCCAGAGAGC	3576
Dp	3642	GTCCCCACGGGGGCAAGGCTGGGGCCACCGGAAGGCTGGGGGGTGCCCAAGGCCAGAGAGC	3701
QY	3577	GTCCGTGCTGAAGAGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAGCAT	3636
Dp	3702	GTCCGTGCTGAAGAGGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCACCAAGCAT	3761
QY	3637	TCCCAAGCAAGGGGTGGCCCTCTGGGACAGGCCCATCAATACCGGGCTTCATCAACCAACGG	3696
Dp	3762	TCCCAAGCAAGGGGTGGCCCTCTGGGACAGGCCCATCAATACCGGGCTTCATCAACCAACGG	3821
QY	3697	CAGCCCAAGCTACGTCCTGTACAAGGGGCAACATCAACAGATTCATCGGCGAGGACAGCC	3756
Dp	3822	CAGCCCAAGCTACGTCCTGTACAAGGGGCAACATCAACAGATTCATCGGCGAGGACAGCC	3881
QY	3757	GAGTTCCTTGGACCGCGCGGCCGGAGAGACAGCTCTGCCAAGGCCACGTCACTCAAGAGG	3816
Dp	3882	GAGTTCCTTGGACCGCGCGGCCGGAGAGACAGCTCTGCCAAGGCCACGTCACTCAAGAGG	3941
QY	3817	CAAGAAGGGGCAACGTCCTGTCTATGAAGGTGGACATGCTGTGACCCAGTGTCTCAAGGA	3876
Dp	3942	CAAGAAGGGGCAACGTCCTGTCTATGAAGGTGGACATGCTGTGACCCAGTGTCTCAAGGA	4001
QY	3877	GGAACGGCAAGAGCAAGCTCAGAGACCCCCCATAGAGACGGCCGCCCAAGCGCAACTATGA	3936
Dp	4002	GGAACGGCAAGAGCAAGCTCAGAGACCCCCCATAGAGACGGCCGCCCAAGCGCAACTATGA	4061
QY	3937	CATGATGGAAGGGCCCGCTGTGGGCAAGACCATCTCTCAGCCAGCATTCGAAGGTCTCATGGG	3996
Dp	4062	CATGATGGAAGGGCCCGCTGTGGGCAAGACCATCTCTCAGCCAGCATTCGAAGGTCTCATGGG	4121
QY	3997	CCGTGGCCATCCCGCGGAGGACACAGCCCCCAACACTCTCAAAAGGAGCAACCAATCCG	4056
Dp	4122	CCGTGGCCATCCCGCGGAGGACACAGCCCCCAACACTCTCAAAAGGAGCAACCAATCCG	4181
QY	4057	CGGCTCATCAACAAGGATTCCTCGGCTCTACGTGAGGCAACGAGAGCTACCTGGC	4116
Dp	4182	CGGCTCATCAACAAGGATTCCTCGGCTCTACGTGAGGCAACGAGAGCTACCTGGC	4241
QY	4117	TCGGAGGCGCAAGCTCCTTAAAGCGGGAAGGACGCGCTCGGCCCAACGCGCTCAAGGGAA	4176
Dp	4242	TCGGAGGCGCAAGCTCCTTAAAGCGGGAAGGACGCGCTCGGCCCAACGCGCTCAAGGGAA	4301
QY	4177	CTGTACCGAGGCTTACAGAGCGCAAGGCCCTTGAGGCTTGAAGCTTGAAGCGGCGCCATGA	4236
Dp	4302	CTGTACCGAGGCTTACAGAGCGCAAGGCCCTTGAGGCTTGAAGCTTGAAGCGGCGCCATGA	4361

QY	4237	GGGCTGTGTGGCCACCGTGGAAAGAGGGGGGCGCTCCATCCATGAGATCCCGCGCAGGA	4236
Db	4362	GGGGCTGTGTGGCCACCGTGGAAAGAGGGGGGCGCTCCATCCATGAGATCCCGCGCAGGA	4421
QY	4297	GCTCGGGACACGGCCGGAGCTGGCCCTGGGCGCGCGGCGGTCAAGAGGGGCTCATAC	4356
Db	4422	GCTGGGGACACGGCCGGAGCTGGCCCTGGGCGCGCGGTCAAGAGGGGCTCATAC	4401
QY	4357	GCAGGGCACCCCGGCTCAAGTACGACACCGGCGCGTCCACACTGGCTTCAAAAAGCAGA	4416
Db	4482	GCAGGGCACCCCGGCTCAAGTACGACACCGGCGCGTCCACACTGGCTTCAAAAAGCAGA	4511
QY	4417	CGTACGCTCCCTCATTCGGCAGCCCCCGGCGCGGACGTTCCACCCGTGCACCCGCTGATGT	4478
Db	4542	CGTACGCTCCCTCATTCGGCAGCCCCCGGCGCGGACGTTCCACCCGTGACCCGCTGATGT	4601
QY	4477	GATGGCGGACGGCCGGGGCATCTGAAACGTGCGCTGATAGAGGAGCTGAAAGACCGGGC	4536
Db	4602	GATGGCGGACGGCCGGGGCATCTGAAACGTGCGCTGATAGAGGAGCTGAAAGACCGGGC	4661
QY	4537	AGGACCGCGCAGCAGCTCGGGGGGCTCATTTGCGCGGCGCGCCCGGTCAATTGGCTGA	4596
Db	4662	AGGACCGCGCAGCAGCTCGGGGGGCTCATTTGCGCGGCGCGCCCGGTCAATTGGCTGA	4721
QY	4597	GCTGGGTAAAGCCGGCGGAGAGCCCCCTGAGCCTATAGAGACCAAGGCAACCTTTGCGG	4656
Db	4722	GCTGGGTAAAGCCGGCGGAGAGCCCCCTGAGCCTATAGAGACCAAGGCAACCTTTGCGG	4781
QY	4657	CCACCTCCCAAGAGTTTCGCCGTGACCATGCGGGAGCCCACGCGCGCTGACAGAGGG	4716
Db	4782	CCACCTCCCAAGAGTTTCGCCGTGACCATGCGGGAGCCCACGCGCGCTGACAGAGGG	4811
QY	4717	CAGCCTTTCCGTCAGCAGAGGCACTCCAGAACGTAAGCTGACAGCCTCGTGAGAT	4776
Db	4842	CAGCCTTTCCGTCAGCAGAGGCACTCCAGAACGTAAGCTGACAGCCTCGTGAGAT	4901
QY	4777	CGCCAAAGTCCCGCGCAGCAGCCGTGCCCGAGCACCAACCAACCCCATTTGCGCCTTGA	4836
Db	4902	CGCCAAAGTCCCGCGCAGCAGCCGTGCCCGAGCACCAACCAACCCCATTTGCGCCTTGA	4961
QY	4837	GCACCTGTCTTGGGGCGTGAAGTGGCGTGAACCTGTATGACAGCCACATCCCCCTGGCCTT	4896
Db	4962	GCACCTGTCTTGGGGCGTGAAGTGGCGTGAACCTGTATGACAGCCACATCCCCCTGGCCTT	5021
QY	4897	CGACCCCAACCTCCATTAACCCGCGGCATCCGCTGGAAGGAGCGGCTGCTTACCTGCC	4956
Db	5022	CGACCCCAACCTCCATTAACCCGCGGCATCCGCTGGAAGGAGCGGCTGCTTACCTGCC	5081
QY	4957	CCGACACTGTGGCCCCCAACCCCACTTACCGGCACTTACCCTTACCTCATCCGCGG	5016
Db	5082	CCGACACTGTGGCCCCCAACCCCACTTACCGGCACTTACCCTTACCTCATCCGCGG	5141
QY	5017	CTTACCCCGACACGGCGGGCGCTGAGAAACGGGACACATCATCAATGACTCATCCTC	5076
Db	5142	CTTACCCCGACACGGCGGGCGCTGAGAAACGGGACACATCATCAATGACTCATCCTC	5201
QY	5077	GCAGCAATAGACCAACAACAGGGCACCGGCACGATGGCCCAAGAGCTAATATGCTGAGGGG	5136
Db	5202	GCAGCAATAGACCAACAACAGGGCACCGGCACGATGGCCCAAGAGCTAATATGCTGAGGGG	5261
QY	5137	CCTTTCGCCCGCAGAGTCTGTGCTGGGCACTCAACTACGCTGGGGTCCCGAGGACATAT	5196
Db	5262	CCTTTCGCCCGCAGAGTCTGTGCTGGGCACTCAACTACGCTGGGGTCCCGAGGACATAT	5321
QY	5197	CGACCTGTCCCAAGTGCACACTTGTCTGTGTCTGTGCCCCCGACACAGGACCCCAAC	5256
Db	5322	CGACCTGTCCCAAGTGCACACTTGTGTGTGTGTGCCCCCGACACAGGACCCCAAC	5381
QY	5257	CACCGCAATGACCGCGCTTGTCTTACCTTCCACCGCGCCCAAGCCTTTCAGAGCCGCA	5316
Db	5382	CACCGCAATGACCGCGCTTGTCTTACCTTCCACCGCGCCCAAGCCTTTCAGAGCCGCA	5441
QY	5317	CAGCAGCTCCCACTCTCCCAAGAGGTCCCAACACATTTGACAAACCAACACACAGTC	5376

Db 5442 CAGAGCTCCCACTCTCCCAAGAGGTCAACACTTGACAAAACCAACCAACGTC 5501  
QY 5377 CTGTCGAGCGGAGAGAGACCGGGATCGAGAGCGGACCGGGATCGGGAGAAA 5436  
Db 5502 CTCGTCGAGCGGAGAGAGACCGGGATCGAGAGCGGACCGGGATCGGGAGAAA 5561  
QY 5437 GTCCATCTCTGACGTTCACCAAGACGGTGAAGACCGACCTCATCTGAGAC 5496  
Db 5562 GTCCATCTCTGACGTTCACCAAGACGGTGAAGACCGACCTCATCTGAGAC 5621  
QY 5497 GAGAGAGCGGAG 5556  
Db 5622 GAG 5681  
QY 5557 CTCCCACTCCCACTGCGCACGACGACGACGACGACGACGACGACGACGAC 5616  
Db 5682 CTCCCACTCCCACTGCGCACGACGACGACGACGACGACGACGACGACGAC 5741  
QY 5617 GAG 5676  
Db 5742 GAG 5801  
QY 5677 GAG 5736  
Db 5802 GAG 5861  
QY 5737 ATTCCCACTGCGCACGACGACGACGACGACGACGACGACGACGACGAC 5796  
Db 5862 ATTCCCACTGCGCACGACGACGACGACGACGACGACGACGACGACGAC 5921  
QY 5797 CATGAG 5856  
Db 5922 CATGAG 5981  
QY 5857 AGCAG 5916  
Db 5982 AGCAG 6041  
QY 5917 CTCCTCCCACTGCGCACGACGACGACGACGACGACGACGACGACGACGAC 5976  
Db 6042 CTCCTCCCACTGCGCACGACGACGACGACGACGACGACGACGACGACGAC 6101  
QY 5977 CACCATGCGCGCACGACGACGACGACGACGACGACGACGACGACGACGAC 6036  
Db 6102 CACCATGCGCGCACGACGACGACGACGACGACGACGACGACGACGACGAC 6161  
QY 6037 GCGCGCACGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 6096  
Db 6162 GCGCGCACGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 6221  
QY 6097 CATCCAG 6156  
Db 6222 CATCCAG 6281  
QY 6157 GGTGAG 6216  
Db 6282 GGTGAG 6341  
QY 6217 GCACTTGAAG 6276  
Db 6342 GCACTTGAAG 6401  
QY 6277 CCCGCTGAAG 6336  
Db 6402 CCCGCTGAAG 6461  
QY 6337 CCAAGCTTGTCTCAAGCGCGCTGTCTCAAGCGCGCTGTCTCAAGCGCGCTGT 6396  
Db 6462 CCAAGCTTGTCTCAAGCGCGCTGTCTCAAGCGCGCTGTCTCAAGCGCGCTGT 6521  
QY 6397 GGTTCACCTGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6456

Db 6522 GGTTCACCTGTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6581  
QY 6457 ACAAG 6516  
Db 6582 ACAAG 6641  
QY 6517 GGTCTTGAAG 6576  
Db 6642 GGTCTTGAAG 6701  
QY 6577 CCGGCGCGGTGCGCTCCCGCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6636  
Db 6702 CCGGCGCGGTGCGCTCCCGCACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6761  
QY 6637 GTGCGTCTTGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6696  
Db 6762 GTGCGTCTTGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6821  
QY 6697 GAG 6756  
Db 6822 GAG 6881  
QY 6757 GAG 6816  
Db 6882 GAG 6941  
QY 6817 CAG 6876  
Db 6942 CAG 7001  
QY 6877 GAG 6936  
Db 7002 GAG 7061  
QY 6937 GATCTTCAAT 6996  
Db 7062 GATCTTCAAT 7121  
QY 6997 GCAG 7056  
Db 7122 GCAG 7181  
QY 7057 TAAATATGACAG 7116  
Db 7182 TAAATATGACAG 7241  
QY 7117 TGCCAG 7176  
Db 7242 TGCCAG 7301  
QY 7177 CACACTCACTTCGACAG 7236  
Db 7302 CACACTCACTTCGACAG 7361  
QY 7237 AAAAGCAAAGTCCCGCGCGCTGCGCATCTGCGGAGAGAGAGAGAGAGAGAGAGAG 7296  
Db 7362 AAAAGCAAAGTCCCGCGCGCTGCGCATCTGCGGAGAGAGAGAGAGAGAGAGAGAG 7421  
QY 7297 AGTGACCTGAG 7356  
Db 7422 AGTGACCTGAG 7481  
QY 7357 CAGAGCTTGTTCGAG 7416  
Db 7482 CAGAGCTTGTTCGAG 7541  
QY 7417 GCGGAGTGTCAATGCTTCCCAAG 7476  
Db 7542 GCGGAGTGTCAATGCTTCCCAAG 7601  
QY 7477 TGAGCCCAAG 7536  
Db 7602 TGAGCCCAAG 7661



QY	7537	ACTCTCCGACGACGAGTACTCAGAAACAAGGGGCGGGGGGGGGCGGGCGGTBUBERSPERLUG	7596
Db	7662	ACTCTCCGACGACGAGTACTCAGAAACAAGGGGCGGGGGGGGGCGGGCGG-----TG	7711
QY	7597	TCAGGTCCGACGACGACCAAGAAACGGCCCTGCAGAGACGGGGCGGCTGCCGACTCCCC	7656
Db	7712	TCAGGTCCGACGACGACCAAGAAACGGCCCTGCAGAGACGGGGCGGCTGCCGACTCCCC	7711
QY	7657	AACCAAGGAAGAGACCCCTGAGTCCGCTCGGCGCTCCATCCATCTGTCCGTCCAGAGCG	7716
Db	7772	AACCAAGGAAGAGACCCCTGAGTCCGCTCGGCGCTCCATCCATCTGTCCAGAGCG	7831
QY	7717	GCATCCTTGGCTGTCTAAAGCCTTAACTAAAGCTCCCGCCCGGGCTGGCCTGTGCAGA	7776
Db	7832	GCATCCTTGGCTGTCTAAAGCCTTAACTAAAGCTCCCGCCCGGGCTGGCCTGTGTGCAGA	7891
QY	7777	CCTTACTCAGGGAGTGTTTACTCGGTGCTCGGAGAGGGAGGGAGAGGGCGCGGGAGGG	7836
Db	7892	CCTTACTCAGGGAGTGTTTACTCGGTGCTCGGAGAGGGAGGGAGGGCGCGGGAGGG	7951
QY	7837	GCAAGGGCAGGGGTGTAGGCAAGCAACAGCGGGCAGGGGGGCGCAGGGACCCAAAGAG	7896
Db	7952	GCAAGGGCAGGGGTGTAGGCAAGCAACAGCGGGCAGGGGGGCGCAGGGACCCAAAGAG	8011
QY	7897	GATGACCAACGACCTCCACAGCCCACTGCTCCCGGAATGATTTTGAACCAAGCTTAA	7956
Db	8012	GATGACCAACGACCTCCACAGCCCACTGCTCCCGGAATGATTTTGAACCAAGCTTAA	8071
QY	7957	CTGAGCTTGCAGCCCCCGCGCCCTCCCTCGGCTCCATCCCGCTTAGCGCTTGCAGAG	8016
Db	8072	CTGAGCTTGCAGCCCCCGCGCCCTCCCTCGGCTCCATCCCGCTTAGCGCTTGCAGAG	8131
QY	8017	ATGAGACGACGAGCCCTGTGCAGCCCCCAGATGAGCTCGTTCCGGTCCGCCACAGACTGCCCA	8076
Db	8132	ATGAGACGACGAGCCCTGTGTCAAGCCCCCAGATGAGCTCGTTCCGGTCCGCCACAGACTGCCCA	8191
QY	8077	GCCAAACGAGATTGCTGGAACCAAGTCAAGCCAGAGTGGGCGGACAAAGGGCCAGTGTGC	8136
Db	8192	GCCAAACGAGATTGCTGGAACCAAGTCAAGCCAGAGTGGGCGGACAAAGGGCCAGTGTGC	8251
QY	8137	GCCGTGGGGGGGACGGATGTCGCGAGACTGGACTGTTTTTTTTCACACATCGTTCGCCGAG	8196
Db	8252	GCCGTGGGGGGGACGGATGTCGCGAGACTGGACTGTTTTTTTTCACACATCGTTCGCCGAG	8311
QY	8197	CGGTGGGAGGAAAGGACAGATGTAATGANTGTGTGTTTTCAGGGTATATTTTGTGATAC	8256
Db	8312	CGGTGGGAGGAAAGGACAGATGTAATGANTGTGTGTTTTCAGGGTATATTTTGTGATAC	8371
QY	8257	CTTCAATGAAATTAATTCAAGATGTTTACGCAAGAGAGCTTACCCAGTATTAATCTGCTGC	8316
Db	8372	CTTCAATGAAATTAATTCAAGATGTTTACGCAAGAGAGCTTACCCAGTATTAATCTGCTGC	8431
QY	8317	TGTGCTTTTGAATCTCTGCTTACCGTTCAAGAGGGCGTGTGCAGGGCCGACAGTCGGGACCC	8376
Db	8432	TGTGCTTTTGAATCTCTGCTTACCGTTCAAGAGGGCGTGTGTGCAGGGCCGACAGTCGGGACCC	8491
QY	8377	CATCACTTCGACGAGCAAGAGGGGGCGGGGACTGCTGTCAAGCCCGCGCTGTGTCTCTCTC	8436
Db	8492	CATCACTTCGACGAGCAAGAGGGGGCGGGGACTGCTGTCAAGCCCGCGCTGTGTCTCTCTC	8551
QY	8437	CCCTCCCTTCCCTTGGGCGGAATGAAATTGCAATCCGATTTCTGTGGCGCGCATTTTGGGCGAGG	8496
Db	8552	CCCTCCCTTCCCTTGGGCGGAATGAAATTGCAATCCGATTTCTGTGGCGCGCATTTTGGGCGAGG	8611
QY	8497	TGTGTGTATTTCTGTCAATTTACACAGCTGTGTTCTAATTTAAAGAGGAATTAATCTCAAA	8556
Db	8612	TGTGTGTATTTCTGTCAATTTACACAGCTGTGTTCTAATTTAAAGAGGAATTAATCTCAAA	8671
QY	8557	AAAAAAAAAAAAAAAA 8571	
Db	8672	AAAAAAAAAAAAAAAA 8686	

XX	ADL12577	standard; cDNA; 9053 BP.
XX	ADL12577;	
XX	06-MAY-2004	(first entry)
XX	Human steroid-induced C3A liver cell cDNA #306.	
XX	88; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;	
XX	steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.	
XX	Homo sapiens.	
XX	US6673549-B1.	
XX	06-JAN-2004.	
XX	12-OCT-2001; 2001US-00976594.	
XX	12-OCT-2000; 2000US-0240409P.	
XX	(INCY-) INCYTE CORP.	
XX	Furness LM, Buchbinder JL;	
XX	WPI; 2004-068610/07.	
PT	Combination useful for preparing a composition for treating liver	
PT	disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,	
PT	comprises cDNAs that are differentially expressed in response to steroid	
PT	treatment.	
PS	Claim 1; SEQ ID NO 306; 141pp; English.	
XX	The invention relates to a combination comprising cDNAs that are	
CC	differentially expressed in response to steroid treatment. Also included	
CC	are the following: a high throughput method for using a cDNA to detect	
CC	differential expression of nucleic acids in a sample; and a high	
CC	throughput method of screening molecules or compounds to identify a	
CC	cDNA that specifically binds a cDNA. The sample is from a subject with	
CC	Wilson disease and comparison of a standard defines a stage of that	
CC	disease. The high throughput method of screening molecules or compounds	
CC	to identify a ligand that specifically binds a cDNA comprises: combining	
CC	the combination with molecules or compounds under conditions to allow	
CC	specific binding; and detecting specific binding between each cDNA and at	
CC	least one molecule or compound. The molecules or compounds are regulatory	
CC	proteins. The combination is useful for preparing a composition for	
CC	treating liver disorders associated with steroid therapy, e.g., cirrhosis	
CC	or hepatitis. The present sequence represents a human cDNA which is	
CC	differentially expressed in steroid-induced C3A liver cells. Note: The	
CC	sequence data for this patent did not form part of the printed	
CC	specification but was obtained in electronic format directly from USPRO	
CC	at seqdata.uspro.gov/sequence.html.	
SO	Sequence 9053 BP; 1904 A; 3171 C; 2722 G; 1254 T; 0 U; 2 Other;	
QY	Query Match	96.1%; Score 8226.8; DB 12; Length 9053;
D5	Best Local Similarity	98.3%; Pred. No. 0;
D5	Matches 8451, Conservative %	2; Mismatches 42; Indels 103; Gaps 10;
OY	1 CATGTCGGGGCTTCCACAGACTGTGGAGCAGACGTGAGGGGCCACTGAGCCCCGCTATCCC	60
D5		
D5	CATTCTGGAGATCCACACAGCGCTGTGGACAGACTGTGAAGGGCCACTGAGCCCCGCTATACC	578
OY	61 GCCCACAGAGCTTTCTTACCAATGCAGATGCGCCGAGACGACACGAGTCTGGGCTCTT	120
D5	579 GCCCACAGAGCTTTCTTACCAATGCAGATGCGCCGAGACGACACGAGTCTGGGCTCTT	638
OY	121 GAAGTACACGACCACTTCGCGCATATGCTTCCACCTGTGCGCGGGCTTCATCATCCA	180
D5		

Dd 639 GGAGTACCAACAACAACCTCCCGCACTATGCTCCACCTGTGCGCCCGGCTCATCATCCA 698  
Qy 181 GCCCAGCGCGCGAGAGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGAAATGAACGGTCCA 240  
Dd 699 GCCCAAGCGGGGAGAGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGAAATGAACGGTCCA 758  
Qy 241 GGAGCTCCACTCTGCGGCGCAAGATCCCACTCATCTGTGCCGAGCTGGGGAAATCAGAGAT 300  
Dd 759 GGAGCTCCACTCTGCGGCGCAAGATCCCACTCATCTGTGCCGAGCTGGGGAAATCAGAGAT 818  
Qy 301 GGAGTTCAATTGAAAGCAAGCGGCTCTGAGCTGAGCTGCGCTGAGACCCCTGCTGAGAC 360  
Dd 819 GGAGTTCAATTGAAAGCAAGCGGCTCTGAGCTGAGCTGCGCTGAGACCCCTGCTGAGAC 878  
Qy 361 GTCAACCCCTGTGTGCGCAACCGGCGCAAGCTGTGAAGACTTCAACCAAGAACCGTAG 420  
Dd 879 GTCAACCCCTGTGTGCGCAACCGGCGCAAGCTGTGAAGACTTCAACCAAGAACCGTAG 938  
Qy 421 CTTGACGGGCAAGCTTGAAGCCGGTGTCTTCCCGCAGCCCGCACTGACCTTGAGCT 480  
Dd 939 CTTGACGGGCAAGCTTGAAGCCGGTGTCTTCCCGCAGCCCGCACTGACCTTGAGCT 998  
Qy 481 GGAGCTGGTGGCGGCAACCGGCTGTCCAAAGGAGAGCTGATCCAGAACATGGAACGGGTGA 540  
Dd 999 GGAGCTGGTGGCGGCAACCGGCTGTCCAAAGGAGAGCTGATCCAGAACATGGAACGGGTGA 1058  
Qy 541 CCGAGAGATCAACCATGTGTAGAGCAGCAGATCTTAACTGAAAGAAAGCAGCAACAGCT 600  
Dd 1059 CCGAGAGATCAACCATGTGTAGAGCAGCAGATCTTAACTGAAAGAAAGCAGCAACAGCT 1118  
Qy 601 GGAGAGAGAGGCTGCGCAAGCCGCGCGAGCTTGAAGGCCCTGTGTCAACCGCGCCATCGA 660  
Dd 1119 GGAGAGAGAGGCTGCGCAAGCCGCGCGAGCTTGAAGGCCCTGTGTCAACCGCGCCATCGA 1178  
Qy 661 GTTGAAGACCGGAGCGTGTGTCAGATCATCTACGACGAGAAACCGGAAGAAAGCTGAAGC 720  
Dd 1179 GTTGAAGACCGGAGCGTGTGTCAGATCATCTACGACGAGAAACCGGAAGAAAGCTGAAGC 1238  
Qy 721 TGCAATCGGATTTCTGAAGAGCTTGGGCGCCGAGGTGAGCTGCGCTGTACACACAGCC 780  
Dd 1239 TGCAATCGGATTTCTGAAGAGCTTGGGCGCCGAGGTGAGCTGCGCTGTACACAGCC 1298  
Qy 781 CTCGCAACACCGGCGAGTATCATGAGAACATCAAAATAAACAGCGGATGCGGAAGAGCT 840  
Dd 1299 CTCGCAACACCGGCGAGTATCATGAGAACATCAAAATAAACAGCGGATGCGGAAGAGCT 1358  
Qy 841 AATCTTGATCTTCAAGAGGAGGATCAAGCTGGAAGCAATGGAAGCAGAAATTTCTGCCA 900  
Dd 1359 AATCTTGATCTTCAAGAGGAGGATCAAGCTGGAAGCAATGGAAGCAGAAATTTCTGCCA 1418  
Qy 901 GCGCTATGACCAAGCTCATGAGAGGCTTGGAAAAAAAGTGAAGCCATCGAAAAACAACC 960  
Dd 1419 GCGCTATGACCAAGCTCATGAGAGGCTTGGAAAAAAAGTGAAGCCATCGAAAAACAACC 1478  
Qy 961 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGCAATCTACGAAAAACAGTTCTCTGAGAT 1020  
Dd 1479 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGCAATCTACGAAAAACAGTTCTCTGAGAT 1538  
Qy 1021 CCGGAAGACGCGGAGCTGCAAGAGCGCATGCAAGAGGAGGAGCGGCGGAGCGGAGTAG 1080  
Dd 1539 CCGGAAGACGCGGAGCTGCAAGAGCGCATGCAAGAGGAGGAGCGGCGGAGCGGAGTAG 1598  
Qy 1081 GCTGTTCATGTGTGCGCGCGCGAGCAGCAGCAGAGGTTCAGAGATCATGATGAGCTCTC 1140  
Dd 1599 GCTGTTCATGTGTGCGCGCGCGAGCAGCAGCAGAGGTTCAGAGATCATGATGAGCTCTC 1658  
Qy 1141 AGAGCAGAGAAACTTGAAGAGCAATGCGCAAGCTGTGCTGCTGCTGCTGCTGCTGTA 1200  
Dd 1659 AGAGCAGAGAAACTTGAAGAGCAATGCGCAAGCTGTGCTGCTGCTGCTGCTGTA 1718  
Qy 1201 CGAGCCTGACCGAGCGGATGAAGTTCATCAACATGAAACGGGCTTATGCGCGCAACCT 1260  
Dd 1719 CGAGCCTGACCGAGCGGATGAAGTTCATCAACATGAAACGGGCTTATGCGCGCAACCT 1778

Qy 1261 GAAAGTGTACAAAGACCGCCAGGTATGAACATGTGAAGTGAAGCAGAGAAAGAACCTT 1320  
Dd 1779 GAAAGTGTACAAAGACCGCCAGGTATGAACATGTGAAGTGAAGCAGAGAAAGAACCTT 1838  
Qy 1321 CCGGAGAAAGTTTCATGACGATCCCAAGAACTTTGCGCTGATTCGATCTTCTTGAAG 1380  
Dd 1839 CCGGAGAAAGTTTCATGACGATCCCAAGAACTTTGCGCTGATTCGATCTTCTTGAAG 1898  
Qy 1381 GAAAGCAGTGGCTGAGTGGCTCTCTATTAATTAATTAATTAATTAATTAATTAATTA 1440  
Dd 1899 GAAAGCAGTGGCTGAGTGGCTCTCTATTAATTAATTAATTAATTAATTAATTAATTA 1958  
Qy 1441 GAGCTGTGAGACGAGACTATCGGCGCGCGGCAAGAGCCAGCAACA--CAGCA 1497  
Dd 1959 GAGCTGTGAGACGAGACTATCGGCGCGCGGCAAGAGCCAGCAACACAGCA 2018  
Qy 1498 GCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1557  
Dd 2019 GCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2078  
Qy 1558 GAAAGATGAG 1617  
Dd 2079 GAAAGATGAG 2138  
Qy 1618 GAAACGACAGAGAGACCTCTTCAAGAGAGACAGACGACCTTCAAGGAGAGACAGCA 1677  
Dd 2139 GAAACGACAGAGAGACCTCTTCAAGAGAGACAGACGACCTTCAAGGAGAGACAGCA 2198  
Qy 1678 CGAGAGAGAGGCTGTGCGCTCAAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 1737  
Dd 2199 CGAGAGAGAGGCTGTGCGCTCAAGAGCGCGGAGAGAGAGAGAGAGAGAGAGAGAG 2258  
Qy 1738 AGGCGGCATCAACCGGCTCAATGAGCTAATGAGAGCCAAACGAGAGAGAGAGAGAGAG 1797  
Dd 2259 AGGCGGCATCAACCGGCTCAATGAGCTAATGAGAGCCAAACGAGAGAGAGAGAGAGAG 2318  
Qy 1798 GCAGAGCGCGGAGCTGAGCTCCATGAGAGCTGATGAGAGTTCTGCTGAGACAGAGAGAG 1857  
Dd 2319 GCAGAGCGCGGAGCTGAGCTCCATGAGAGCTGATGAGAGTTCTGCTGAGAGAGAGAGAG 2378  
Qy 1858 AATGGAACACGCAAGAAAGGTCTCTTGAAGACAGCGGCGGCAATGAGTGGCGCATGCGCG 1917  
Dd 2379 AATGGAACACGCAAGAAAGGTCTCTTGAAGACAGCGGCGGCAATGAGTGGCGCATGCGCG 2438  
Qy 1918 GATGCTGGCTTCAAGACCTGCTGAGTGAAGCTTCACTTCACTTCACTTCACTTCACTTCA 1977  
Dd 2439 GATGCTGGCTTCAAGACCTGCTGAGTGAAGCTTCACTTCACTTCACTTCACTTCACTTCA 2498  
Qy 1978 GCAGAACCTCGATGAGATCTTGCAGCAGCACAAGCTGGAAGTGAAGAGAGAGAGAGAG 2037  
Dd 2499 GCAGAACCTCGATGAGATCTTGCAGCAGCACAAGCTGGAAGTGAAGAGAGAGAGAGAGAG 2558  
Qy 2038 GCGAG 2096  
Dd 2559 GCGAG 2618  
Qy 2097 TGATGAGAGATGAGAGATGAG 2156  
Dd 2619 TGATGAGAGATGAGAGATGAG 2678  
Qy 2157 AGGAGCTGAAGCCTTACATGCTCTTGGGAAATGAGGTGCCAGAGAGGGAATGCACTGCGC 2216  
Dd 2679 AGGAGCTGA----- 2688  
Qy 2217 CAGCCACTGTCAACAACGCTCAAGACACCGAGAGATCCCTCTTCTCACTGAGGCGG 2276  
Dd 2689 -AGCCACTGTCAACAACGCTCAAGACACCGAGAGATCCCTCTTCTCACTGAGGCGG 2747  
Qy 2277 CCAAGGACACAGGGGAGAGTGGGCGCAGCGCCAGCCAGCCCTGGGCGCGAGC-GGGCGCA 2335  
Dd 2748 CCAAGGACACAGGGGAGAGTGGGCGCAGCGCCAGCCAGCCCTGGGCGCGAGCGGGCGCA 2807

2336 CCCCAAGCCCAACCCCAACGAGAGCATCCCGGCCCCCATTTGAGCCCAACCCG 2395  
2808 CCCCCAGGGCAACCCACCAACGAGAGATCCCGCCCCCACTAGGCCCAACCCG 2867  
2396 GCGCTTGAAAGCAACGAGCCCTTAAGCCCCCAACGACACCCCATGCGCTTGCACT 2455  
2868 GCGCTTGAAAGCAACGAGCCCTTAAGCCCCCAACGACACCCCATGCGCTTGCACT 2927  
2456 CCGCTGTGGTCCCAAGAGAGAGAGAGAGACCGCAACGCGCCCAAGTGGAG 2515  
2928 CCGCTGTGGTCCCAAGAGAGAGAGAGAGACCGCAACGCGCCCAAGTGGAG 2987  
2516 GAGGGAG 2575  
2988 GAGGGAG 3047  
2576 GAG 2635  
3048 GAG 3107  
2636 GAG 2695  
3108 GAG 3167  
2696 GAG 2755  
3168 GAG 3227  
2756 GCTACCTGAGTGCAGAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2815  
3228 GCTACCTGAGTGCAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3287  
2816 TCCCAAG 2875  
3288 TCCCAAG 3347  
2876 AAGCACTGGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2935  
3348 AAGCACTGGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3404  
2936 ACCAAAGTCCATGAG 2994  
3405 ACCAAAGTCCATGAG 3464  
2995 CCCAAG 3054  
3465 CCCAAG 3524  
3055 CCCCGAG 3100  
3525 CCCCGAG 3584  
3101 GCGCTTGAG 3150  
3585 GTTCTTCCAG 3644  
3151 TTCCGAG 3210  
3645 TTCCGAG 3704  
3211 GAG 3270  
3705 GAG 3764  
3271 GAG 3330  
3765 GAG 3824  
3331 GAG 3390  
3825 GAG 3884  
3391 GAG 3450

3885 GGTTCAGTCTCAAGTCCCGTACTCAGAGATGCAAGAGCCCGGAGGCGCTGTCAACAT 3944  
3451 GGGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3510  
3945 GGGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4004  
3511 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3570  
4005 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4064  
3571 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3630  
4065 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4124  
3631 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3690  
4125 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4184  
3691 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3750  
4185 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4244  
3751 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3810  
4245 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4304  
3811 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3870  
4305 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4364  
3871 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3930  
4365 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4424  
3931 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3990  
4425 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4484  
3991 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4050  
4485 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4544  
4051 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4110  
4545 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4604  
4111 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4170  
4605 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4664  
4171 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4230  
4665 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4724  
4231 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4290  
4725 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4784  
4785 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4844  
4351 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4410  
4845 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4904  
4411 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4470  
4905 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4964  
4471 GAGAGTGGCCCGCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4530

Db 4965 GAGATGTGATGCGCAGCGCCCGGACCTGGAACGTGCTGTCAAGAGAGGCTTGAAAG 5024  
Qy 4531 CCGGCACGAGGACCGGCAGAGAGCTCGGAGGAGCTCAATTGCGCGGCGCCCGTCAATTGT 4590  
Db 5025 CCGGCACGAGGACCGGCAGAGAGCTCGGAGGAGCTCAATTGCGCGGCGCCCGTCAATTGT 5084  
Qy 4591 GCTGAGAGCTGGGTAAAGCGGCGAGAGCCCTTGACCTATGAGAGCAACGAGGAGCACTTT 4650  
Db 5085 GCTGAGAGCTGGGTAAAGCGGCGAGAGCCCTTGACCTATGAGAGCAACGAGGAGCACTTT 5144  
Qy 4651 TGGCGGCGACCTTCCACGAGAGGTTCCCGCGTGAACATGCGGAGGCCACGCGCGCTTGA 4710  
Db 5145 TCGCGGCGACCTTCCACGAGAGTTCCCGCGTGAACAGCGGAGGCCACGCGCGCTTGA 5204  
Qy 4711 GAGAGGCGACCTTGTGTCCAGAGAGCATCCGAGAGCCGAAGACTGACGTGACGCTTG 4770  
Db 5205 GAGAGGCGACCTTGTGTGTCCAGAGAGCATCCGAGAGCCGAAGACTGACGTGACGCTTG 5264  
Qy 4771 TGAAGATCGCAAGTCCCGCGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4830  
Db 5265 TGAAGATCGCAAGTCCCGCGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5324  
Qy 4831 CTATAGAGACCTGCTTGTGCGGAGGTGAGAGGAGTGAACCTGTATTCGAGAGCAATCCCTT 4890  
Db 5325 CTATAGAGACCTGCTTGTGCGGAGGTGAGAGGAGTGAACCTGTATTCGAGAGCAATCCCTT 5384  
Qy 4891 GGCCTTTCAGACCCACCTCCATATCCCGCGGACATCCCTCTGAGAGGAGCGCTGCTACTA 4950  
Db 5385 GGCCTTTCAGACCCACCTCCATATCCCGCGGACATCCCTCTGAGAGGAGCGCTGCTACTA 5444  
Qy 4951 CTTGCTTCCAGACCTGAGGCGCCCAACCCCACTTACCTGTAACCTGTAACCTGTAACCT 5010  
Db 5445 CTTGCTTCCAGACCTGAGGCGCCCAACCCCACTTACCTGTAACCTGTAACCTGTAACCT 5504  
Qy 5011 CCGGCGCTACCCCGGACAGCGGCGCGCTGAGAGAACGCGGAGACCATATCAATGATCAAT 5070  
Db 5505 CCGGCGCTACCCCGGACAGCGGCGCGCTGAGAGAACGCGGAGACCATATCAATGATCAAT 5564  
Qy 5071 CACCTTCGAGAGAGATGACCAACAACGCGGACCGGACATGAGCCGAGAGAGTATGCT 5130  
Db 5565 CACCTTCGAGAGAGATGACCAACAACGCGGACCGGACATGAGCCGAGAGAGTATGCT 5624  
Qy 5131 GAGGGGCTCTTGGCGCGGAGTCTTGTGAGCACTCAATGCTGCGGAGTCCCGAGG 5190  
Db 5625 GAGGGGCTCTTGGCGCGGAGTCTTGTGAGCACTCAATGCTGCGGAGTCCCGAGG 5684  
Qy 5191 CATCATTCGACCTGTCCCAAGTGCACACCTGAGCTGTGCTGCGGACCGGACCAAGGAC 5250  
Db 5685 CATCATTCGACCTGTCCCAAGTGCACACCTGAGCTGTGCTGCGGACCGGACCAAGGAC 5744  
Qy 5251 CCGAGCCACCGGCATGAGACCGGCTTGCTACTCTCCACCGGCGCCAGCCCTTGAGCAG 5310  
Db 5745 CCGAGCCACCGGCATGAGACCGGCTTGCTACTCTCCACCGGCGCCAGCCCTTGAGCAG 5804  
Qy 5311 CCGGCACAGAGCTTCCCACTCTCTCCAGAGAGTCCAAACAATTGACAAAACAACAC 5370  
Db 5805 CCGGCACAGAGCTTCCCACTCTCTCCAGAGAGTCCAAACAATTGACAAAACAACAC 5864  
Qy 5371 CACGCTCTGTCCGAGCGGAGCGGAGCGGAGATGAGAGCGGAGACCGGAGATCGGAGAG 5430  
Db 5865 CACGCTCTGTCCGAGCGGAGCGGAGCGGAGATGAGAGCGGAGACCGGAGATCGGAGAG 5924  
Qy 5431 GGAAGAGTTCATCTTCATCAGTTCACCAAGAGAGTGAAGCAAGCAATCTGAGAGCTTG 5490  
Db 5925 GGAAGAGTTCATCTTCATCAGTTCACCAAGAGAGTGAAGCAAGCAATCTGAGAGCTTG 5984  
Qy 5491 TACAG 5550  
Db 5985 TACAG 6044  
Qy 5551 CCGCGCTTCCCACTTCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5610  
Db 6045 CCGCGCTTCCCACTTCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6104

Qy 5611 CTTCCAG 5670  
Db 6105 CTTCCAG 6164  
Qy 5671 GAGAGCCGAG 5721  
Db 6165 GAGAGCCGAG 6224  
Qy 5722 TGGCCAG 5781  
Db 6225 TGGCCAG 6284  
Qy 5782 GGTCTACCTTACCTCATGAG 5841  
Db 6285 GGTCTACCTTACCTCATGAG 6344  
Qy 5842 GCGAG 5901  
Db 6345 GCGAG 6404  
Qy 5902 CCGGCTGAG 5961  
Db 6405 CCGGCTGAG 6464  
Qy 5962 TGTCTCTGAG 6021  
Db 6465 TGTCTCTGAG 6524  
Qy 6022 CAGCCCGAG 6081  
Db 6525 CAGCCCGAG 6584  
Qy 6082 AAGTAAACCTTTTTCATTCAG 6141  
Db 6585 AAGTAAACCTTTTTCATTCAG 6644  
Qy 6142 CTACAG 6201  
Db 6645 CTACAG 6704  
Qy 6202 CAG 6261  
Db 6705 CAG 6764  
Qy 6262 GCGCAAG 6321  
Db 6765 GCGCAAG 6824  
Qy 6322 GCGGCTGCTGAG 6381  
Db 6825 GCGGCTGCTGAG 6884  
Qy 6382 AGGTACCAAG 6441  
Db 6885 AGGTACCAAG 6944  
Qy 6442 CACCGGAG 6501  
Db 6945 CACCGGAG 7004  
Qy 6502 TGGGCGAG 6561  
Db 7005 TGGGCGAG 7064  
Qy 6562 CCGGAG 6621  
Db 7065 CCGGAG 7124  
Qy 6622 AGAGCAAAAG 6681  
Db 7125 AGAGCAAAAG 7184

OY	6682	ACCGGAGGGGCAATGACGGAGCCAGGGGCACTCCCGAGAGCTGTGTATCCCGGCTGTATCG	6741
Db	7185	ACCGGAGGGGCAATGACGGAGCCAGGGGCACTCCCGAGAGCTGTGTATCCCGCTGTGTACG	7244
OY	6742	GGATTGGGGAAACAGACGGAGCCAGAGGATGGGCTCCAAAGTCTCCAGGCAACACACGCCA	6801
Db	7245	GGATTGGGGAAACAGACGGAGCCAGAGGATGGGCTCCAAAGTCTCCAGGCAACACACGCCA	7304
OY	6802	GCCGCGACGCTTTCTTTCAGCAAGCTGACCGGAGGCACTCCGCGCATGTGTCAAGTCCAGAA	6861
Db	7305	GCCGCGACGCTTTCTTTCAGCAAGCTGACCGGAGGCACTCCGCGCATGTGTCAAGTCCAGAA	7364
OY	6862	GCAAGAGATTCACAAAGAAAGCTGGAACACCCCAACAACCGGAATGAGGCTGAATTAATATACG	6921
Db	7365	GCAAGAGATTCACAAAGAAAGCTGGAACACCCCAACAACCGGAATGAGGCTGAATTAATATACG	7424
OY	6922	CGAGGCTTGGGACGGAGATCTTCAATATATCCCGCATTCACCGGAAACAGGCTTTATGACCTA	6981
Db	7425	CGAGGCTTGGGACGGAGATCTTCAATATATCCCGCATTCACCGGAAACAGGCTTTATGACCTA	7484
OY	6982	TAGAAAGCAGGCGGCTGCAAGAAACATGCGAGCAACCAATGCGGCTGAGAGGCATATATTAG	7041
Db	7485	TAGAAAGCAGGCGGCTGCAAGAAACATGCGAGCAACCAATGCGGCTGAGAGGCATATATTAG	7544
OY	7042	AAAGGCACTCATGSGGTAAATATATGACCAATGGGAAAGATCCCCCGGCTCAGCGCCATATG	7101
Db	7545	AAAGGCACTCATGSGGTAAATATATGACCAATGGGAAAGATCCCCCGGCTCAGCGCCATATG	7604
OY	7102	TTTTTAACCTCTGAATATGCGAGGCGGCTGCGGCTGTATAGCCATTAACCGGCTGCTGA	7161
Db	7605	TTTTTAACCTCTGAATATGCGAGGCGGCTGCGGCTGTATAGCCATTAACCGGCTGCTGA	7664
OY	7162	CGGACGGAGTGCACACACACTCACTCCGCAAGTGGCGGCGGAGGCAAGCGTCTCTGCG	7221
Db	7665	CGGACGGAGTGCACACACACTCACTCCGCAAGTGGCGGCGGAGGCAAGCGTCTCTGCG	7724
OY	7222	CAGACCCAGCAGCGGAAAGCCAGTCCCCGGGCTCCGGGCTGTGGCATTTGGGAAACGGGCT	7281
Db	7725	CAGACCCAGCAGCGGAAAGCCAGTCCCCGGGCTCCGGGCTGTGGCATTTGGGAAACGGGCT	7784
OY	7282	AACCTCTGTCTCTTCAGTGCACCTGSGAAGGAGACTGGAACCGCGGGAACGGCGCTCACCAA	7341
Db	7785	AACCTCTGTCTCTTCAGTGCACCTGSGAAGGAGACTGGAACCGCGGGAACGGCGCTCACCAA	7844
OY	7342	CCGCGTGTGGAGAGACAGGAGCCCTCGTCCGAGAGTTCCAGCGCACTTCCCTTCAACCCCT	7401
Db	7845	CCGCGTGTGGAGAGACAGGAGCCCTCGTCCGAGAGTTCCAGCGCACTTCCCTTCAACCCCT	7904
OY	7402	GATCATGCGGCTGCGAGGCGGAGTGTATGATGATTCCTCCACCCCAACCGGAGCTCCCGCGGG	7461
Db	7905	GATCATGCGGCTGCGAGGCGGAGTGTATGATGATTCCTCCACCCCAACCGGAGCTCCCGCGGG	7964
OY	7462	CAGCGGAGCCCTCTGCTGTGGCCCCCACCACACAGCTTGGGAGGAGAGCCCAAGCAGTGTCTG	7521
Db	7965	CAGCGGAGCCCTCTGCTGTGGCCCCCACCACACAGCTTGGGAGGAGAGCCCAAGCAGTGTCTG	8024
OY	7522	CTCGCAGTACGAGACACTCTCCGACAGGAGGTGACTCAGAAACAGGAGCGGAGGAGGAGCGG	7581
Db	8025	CTCGCAGTACGAGACACTCTCCGACAGGAGGTGACTCAGAAACAGGAGCGGAGGAGGAGCGG	8084
OY	7582	CGGTGEBURS PERLUGTCAAGTTCACGAGCCACAGAAACGAGCCCTGCAAGAGCGGAGCG	7641
Db	8085	-----GGGATGTCAAGTTCACGAGCCACAGAAACGAGCCCTGCAAGAGCGGAGCGG	8135
OY	7642	GCTGCCGACTCCCCCAACCAAGAAAGAGCCCTGAGATTCGCTGCGCTCTCATATCATCT	7701
Db	8136	GCTGCCGACTCCCCCAACCAAGAAAGAGCCCTGAGATTCGCTGCGCTCTCATATCATCT	8195
OY	7702	GTCCGTCCAGAGCGGAGATTCCTGTGCTGTCTTAAAGCCTTTAAGCTAAGACTCCCGGCGCGG	7761
Db	8196	GTCCGTCCAGAGCGGAGATTCCTGTGCTGTCTTAAAGCCTTTAAGCTAAGACTCCCGGCGCGG	8255
OY	7762	CTGGGCTCTGTGAGACTTTACTCAGAGGAGTGTTAACCTGTGTCTCGGAAAGGAGGGGAA	7821

Db	8256	CTGGCCCTGTGACAGCTTAACTCAGGGAGTCTTAACTCGGTGCTCGGGAAAGGAGGGGAA	8315
Qy	7822	GGGGCCGGGGAGGGGGGCAACGGCAGGCGTGTGGACCAACAACAAGCGGCGCAAGGGCGGCC	7881
Db	8316	GGGGCCGGGGAGGGGGGCAACGGCAGGCGTGTGGACCAACAACAAGCGGCGCGCAAGGGCGGCC	8375
Qy	7882	AGGGAACCCAAAGCAGATGACACACAGCACTTCCAGCCACAGCCCTCCCCCGAAATGCATTG	7941
Db	8376	AGGGAACCCAAAGCAGATGACACAGCACTTCCAGCCACAGCCCTCCCCCGAAATGCATTG	8435
Qy	7942	GAACCAAAAGTCTAAACCTGAGCTCGCAGCCCGCGGCCCTCCCTCGGCTCCCATCCGCT	8001
Db	8436	GAACCAAAAGTCTAAACCTGAGCTCGCAGCCCGCGGCCCTCCCTCGGCTCCCATCCGCT	8495
Qy	8002	TAGGCGCTCTGGACAGATGACAGCAGGCCCTGTCCAGGCCCCCAAGTGCCTCGTTCCGGTCC	8061
Db	8496	TAGGCGCTCTGGACAGATGACAGCAGGCCCTGTCCAGGCCCCCAAGTGCCTCGTTCCGGTCC	8555
Qy	8062	CCACAGACTGCCCCCAGCCCAACAGATGCTGGAAACCAAGTCAGGCCAGAGTGGCGGACA	8121
Db	8556	CCACAGACTGCCCCCAGCCCAACAGATGCTGGAAACCAAGTCAGGCCAGAGTGGCGGACA	8615
Qy	8122	AAAGGGCAGGTGGGGCTGGGGGGGAAACGATGCTCCAGGACCTGGACTGTTTTCAC	8181
Db	8616	AAAGGGCAGGTGGGGCTGGGGGGGAAACGATGCTCCAGGACCTGGACTGTTTTCAC	8675
Qy	8182	ACATCGTTGCGCAGCGGTGGAGAAAGCAAGCAGATGTAATGATGTGTGGTTTACAGG	8241
Db	8676	ACATCGTTGCGCAGCGGTGGAGAAAGCAAGCAGATGTAATGATGTGTGGTTTACAGG	8735
Qy	8242	GTAATATTTTTGATACCTTCAATGAATTAATTACAGATGTTTACGCAAGAAAGACTTACC	8301
Db	8736	GTAATATTTTTGATACCTTCAATGAATTAATTAGATGTTTACGCAAGAAAGACTTACC	8795
Qy	8302	CAGATTAACGCTGCTGCTGCTTTGATCTGCTGCTAACGTTCAAGAGGCGTGGCAGGCC	8361
Db	8796	CAGATTAACGCTGCTGCTGCTTTGATCTGCTGCTAACGTTCAAGAGGCGTGGCAGGCC	8855
Qy	8362	GACAGTCGTGACCCCACTCACTCGCAGAACCAAGGGGGCGGGACCTGCTG-TCAGCCCC	8420
Db	8856	GACAGTCGTGACCCCACTCACTCGCAGAACCAAGGGGGCGGGACCTGCTGCTGACGCC	8915
Qy	8421	CGCTGTGCTCCTCCCTCCCTTCTCTTGGGAGATGAATGATGATGCTGTGTGCG	8480
Db	8916	CGCTGTGCTCCTCCCTCCCTTCTCTTGGGAGATGAATGATGATGCTGTGTGCG	8975
Qy	8481	CGCCATTTGGCAGAGGTGTGTATTTCTGTCAATTACACAGTCGTCTTAATTTAAAGC	8540
Db	8976	CGCCATTTGGCAGAGGTGTGTATTTCTGTCAATTACACAGTCGTCTTAATTTAAAGC	9035
Qy	8541	GAATTATACTCACAAAAA 8558	
Db	9036	GAATTATACTCACAAAAA 9053	
RESULT 12			
ADQ23294/c			
ID	ADQ23294	standard; DNA; 9079 BP.	
XX	AC	ADQ23294;	
XX	DT	26-AUG-2004 (first entry)	
XX	DE	Human soft tissue sarcoma-upregulated DNA - SRO ID 6114.	
XX	KW	soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;	
XX	OS	Homo sapiens.	
XX	FN	WO2004048938-A2.	





QY	1620	ACGACAGAGAAAGACTCTCTCAAGAGAGACAGACGACCTCAGGGAGGACAAACGAC	1679
Db	6939	ACGACAGAGAAAGACTCTCTCAAGAGAGAGACAGACGACCTCAGGGAGGACAAACGAC	6880
QY	1680	AGAAAGAGGCTGTGGCTCTCAAAAGGCCCAAACTGGCCACACGAGCCAGGGAAAGACGAAAG	1739
Db	6879	AGAAAGAGGCTGTGGCTCTCAAAAGGCCCAAACTGGCCACACGAGCCAGGGAAAGACGAAAG	6820
QY	1740	GCCGATCACCCGCTCAATGACTAATGAGGCCCAACAGCGAGGAGGCACTCACCCCGACG	1799
Db	6819	GCCGATCACCCGCTCAATGACTAATGAGGCCCAACAGCGAGGAGGCACTCACCCCGACG	6760
QY	1800	AGAGCGCCGAGCTGGCTCTCATGAGCTGAATGAGAGTTCTTGCTGACAGAAAGAAA	1859
Db	6759	AGAGCGCCGAGCTGGCTCTCATGAGCTGAATGAGAGTTCTTGCTGACAGAAAGAAA	6700
QY	1880	TGGAAACAGCCCAAGAAAGTCTCTCTGGAAACAGGCCGCACTGGCTGGCCATTGGCCCGGA	1919
Db	6699	TGGAAACAGCCCAAGAAAGTCTCTCTGGAAACAGGCCGCACTGGCTGGCCATTGGCCCGGA	6640
QY	1920	TGGTGGGCTCCAAAGCTGTGTGCGAGTGAAGAACTTCTACTTCAACTACAAAGAAAGGC	1979
Db	6639	TGGTGGGCTCCAAAGCTGTGTGCGAGTGAAGAACTTCTACTTCAACTACAAAGAAAGGC	6580
QY	1980	AGAACCTCGATGAGATCTTTCAGCAGCACAAGCTGAAGATGAGAGAGAGAAACGCGC	2039
Db	6579	AGAACCTCGATGAGATCTTTCAGCAGCACAAGCTGAAGATGAGAGAGAGAAACGCGC	6520
QY	2040	GGAAGAAAGAAAGAAAGCG	2099
Db	6519	GGAAGAAAGAAAGAAAGCG	6460
QY	2100	TGAGAGATGAGAGAGATGAGAGGCGTGGCGCGTGAAGCGAAATGAGAGAGAGATGTGTGAGG	2159
Db	6459	TGAGAGATGAGAGAGATGAGAGGCGTGGCGCGTGAAGCGAAATGAGAGAGAGATGTGTGAGG	6400
QY	2160	AGGCTGAAGCCTTACATGCTCTGTGGAAATGAGGTGCCAGAGAGGGAAATGACGTGGCCAG	2219
Db	6399	AGGCTGA-----AGGCTGA-----AG	6391
QY	2220	CCACTGTCAACACAGCTTCAAGCACCGAGAGCATCCTCTTCCTCACTGAGGCCGCA	2279
Db	6390	CCACTGTCAACACAGCTTCAAGCACCGAGAGCATCCTCTTCCTCACTGAGGCCGCA	6331
QY	2280	AGGACACAGGGCAAGATGGGCGCCAAAGCCCCCAAGCCCACTGTGGGGCGGACGGGCAACCC	2339
Db	6330	AGGACACAGGGCAAGATGGGCGCCAAAGCCCCCAAGCCCACTGTGGGGCGGACGGGCAACCC	6271
QY	2340	CAGGCGCACCCACCCACACAGAGAGACATCCGGGGCCCCCATTTGAGGCCCAACCCCGGCT	2399
Db	6270	CAGGCGCACCCACCCACACAGAGAGACATCCGGGGCCCCCATTTGAGGCCCAACCCCGGCT	6211
QY	2400	CTGAAGCCACCGGAGCCCTTACGCCCCCAGCAGACACCCCATCGCCTTTCGACCTTCTC	2459
Db	6210	CTGAAGCCACCTTACGCCCCCAGCAGACACCCCATTTCTCTTCACTTCTCTC	6151
QY	2460	CTGTGTCTCCCCAAGAGAGAAAGAGAGAGACCGCAGACAGCCGCCCTCACTGAGAGAG	2519
Db	6150	CTGTGTCTCCCCAAGAGAGAGAAAGAGAGAGACCGCAGACAGCCGCCCTCACTGAGAGAG	6091
QY	2520	GGGAGAGACAAAGACCCCGCGCGCTGAGAGCTGGCAGTGAACACAGGGAAAGCCGAGG	2579
Db	6090	GGGAGAGACAAAGACCCCGCGCGCTGAGAGCTGGCAGTGAACACAGGGAAAGCCGAGG	6031
QY	2580	AGCCCGTCAAGAGCGAGTGCACGAGAGAAAGCTGAGAGAGGCGCGCCAAAGGACAGGACG	2639
Db	6030	AGCCCGTCAAGAGCGAGTGCACGAGAGAAAGCTGAGAGAGGCGCGCCAAAGGACAGGACG	5971
QY	2640	CGGAGAGCCGCTGAGAGCCACGGCCGAGGGGGGCGTCAAGCAGAGAAAGAAAGAGGGCGGGA	2699
Db	5970	CGGAGAGCCGCTGAGAGCCACGGCCGAGAGGGGCGTCAAGCAGAGAAAGAAAGAGGGCGGGA	5911

QY	2700	GGGGGAGGGGACCACTTCGACCAAGAGCTGGGGGCGCCCCCAAGACAGGCACTTCAGTGGCTA	2753
Db	5910	GGGGGAGGGGACCACTTCGACCAAGAGCTGGGGGCGCCCCCAAGACAGGCACTTCAGTGGCTA	5851
QY	2760	CGTGAAGTGCAGAGGTTGATGAGGCGGAGGCGGAGCAAGAAACGGGCTGCTGCC	2819
Db	5850	CCTGCAAGTGCAGAGGTTGATGAGGCGGAGGCGGAGCAAGAAACGGGCTGCTGCC	5791
QY	2820	CAAGGCCAGGCTCTCTCACCCCGATCTGGCGACCCCCGGGGCCAAATGCTCAACCCCAAGAGC	2879
Db	5790	CAAGGCCAGGCTCTCTCACCCCGATCTGGCGACCCCCGGGGCCAAATGCTCAACCCCAAGAGC	5731
QY	2880	CACCTGACCTTGAAGAGCTTGAAGCAGCGAGGGCTGGCCATCCCGCCCATCCAGGTCAACA	2939
Db	5730	CACCTGACCTTGAAGAGCTTGAAGCAGCGAGGGCTGGCCATCCCGCCCATCCAGGTCAACA	5671
QY	2940	AAGTCCATGAGCCCCCGGGAGGACGAGCTCCACCAACCAAGCTCCCGACGCCCCAC	2999
Db	5670	AAGTCCATGAGCCCCCGGGAGGACGAGCTCCACCAACCAAGCTCCCGACGCCCCAC	5611
QY	3000	CGCCACCGCAAAACCTTGAAGCCGGAGAGCGACGACCCCTCAGACAGCTTGGCAGACGCCCC	3059
Db	5610	CGCCACCGCAAAACCTTGAAGCCGGAGAGCGACGACCCCTCAGACAGCTTGGCAGACGCCCC	5551
QY	3060	GGGGGAGAGAGAGAGGCCCGGACACCCCGCGCGACAAAGAG-----3100	
Db	5550	GGGGGAGAGAGAGAGGCCCGGACACCCCGCGCGACAAAGAGCGTGTGTTCT	5491
QY	3101	-----GCGTTGGCAGCGGAGGCCCAAGAGCTGGCGGAGACCCCGTTGGCTGAGCTTCG	3155
Db	5490	TCGCCAGCTTTGCAAGCCGAGGCCCAAGAGCTGGCGGAGACCCCGTTGGCTGAGCTTCG	5431
QY	3156	GCGTGCCTTTCCCGGTGCCCCCGGTGAGGTATCAAGGCTCTCCCGCATGCCCGGAGC	3215
Db	5430	GCGTGCCTTTCCCGGTGCCCCCGGTGAGGTATCAAGGCTCTCCCGCATGCCCGGAGC	5371
QY	3216	CCTAGGCTTTCTCTAAGCTCACCTGAGTCAACCACTGACCCCTTGAGGCTTCATGACACTG	3275
Db	5370	CCTAGGCTTTCTCTAAGCTCACCTGAGTCAACCACTGACCCCTTGAGGCTTCATGACACTG	5311
QY	3276	CCCGGCGGCTCTGCGCGGCCACCCACATCTCCAACCGGCTCTCCCTCATCTCTCTG	3335
Db	5310	CCCGGCGGCTCTGCGCGGCCACCCACATCTCCAACCGGCTCTCCCTCATCTCTCTG	5251
QY	3336	CCAAAGCAACCCAGAGGCTCTGAGAGGGAATAATGATGTCATCTCCCAAGGAATGTCGGTCC	3395
Db	5250	CCAAAGCAACCCAGAGGCTCTGAGAGGGAATAATGATGTCATCTCCCAAGGAATGTCGGTCC	5191
QY	3396	AGCTTCAAGTCCCGTAAGAGAGTGCACAAAGCCCCGGTGGGCCCTGTCACTATGGAGGC	3455
Db	5190	AGCTTCAAGTCCCGTAAGAGAGTGCACAAAGCCCCGGTGGGCCCTGTCACTATGGAGGC	5131
QY	3456	TGCCCCCTGCCATGAGACCCCAAAAAGCTGGCACCTTTCAAGCGAGTGAAGCAGAGCAGC	3515
Db	5130	TGCCCCCTGCCATGAGACCCCAAAAAGCTGGCACCTTTCAAGCGAGTGAAGCAGAGCAGC	5071
QY	3516	TGTGCCCAACGGGGGCAAGGCTGGGCGCAACCGAGACCTGGGGGTGCCCAACGCCAGGAGG	3575
Db	5070	TGTGCCCAACGGGGGCAAGGCTGGGCGCAACCGAGACCTGGGGGTGCCCAACGCCAGGAGG	5011
QY	3576	CGTCCGTTGCTGAGAGGAGCAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCAACAAAGCA	3635
Db	5010	CGTCCGTTGCTGAGAGGAGCAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCAACAAAGCA	4951
QY	3636	TTTCCAGAGCACGGGTGCGCTTCGAGACAGCGCATCATACCGCGGCTTCATCAACCAACG	3695
Db	4950	TTTCCAGAGCACGGGTGCGCTTCGAGACAGCGCATCATACCGCGGCTTCATCAACCAACG	4891
QY	3696	GCAAGCCAGCTGACGTCCTGTAACAAGGGGCACATACCAAGATCATGGGAGGAGCAGCC	3755
Db	4890	GCAAGCCAGCTGACGTCCTGTAACAAGGGGCACATACCAAGATCATGGGAGGAGCAGCC	4831
QY	3756	CGAGTCCCTTGAACCGGCGCGGGAGGACAGCTGCCCAGAGGGCACGTCTATCAACAG	3815

|||||  
4830 CGAGTCGCTTGGACCGCGCCGCGGAGGACAGCCTGCCAAGGCGCACGTCATCTGACGAG 4771  
3816 GCANAAGGGGCCAGCTTTGTCTTATGAGGGGTGGATGTCTGTGACCCAGTGTCTCAAGG 3875  
4770 GCANAAGGGGCCAGCTTTGTCTTATGAGGGGTGGATGTCTGTGACCCAGTGTCTCAAGG 4711  
3876 AGGACGGAGAGAGCAGCTCAGGAGCCCCCATAGAGCGGCGGCCCAAGCGACACTATG 3935  
4710 AGGACGGAGAGAGCAGCTCAGGAGCCCCCATAGAGCGGCGGCCCAAGCGACACTATG 4651  
3936 ACATGATGAGGGGCCGCTGGGCAAGGCCATCTCTCAGCCAGCATTCGAAAGTCTCATGG 3995  
4650 ACATGATGAGGGGCCGCTGGGCAAGGCCATCTCTCAGCCAGCATTCGAAAGTCTCATGG 4591  
3996 GCCCGGCATCCCGCGGAGCGACACGCCCCACACCTCAAGAGGACGACCATCC 4055  
4590 GCCGTGCCATCCCGCGGAGCGACACGCCCCACACCTCAAGAGGACGACCATTC 4531  
4056 GCGGCTCATCAACAAGGAGATCCCTCGGTCTTACGTGAGGACAGAGAGACTACCTGC 4115  
4530 GCGGCTCATCAACAAGGAGATCCCTCGGTCTTACGTGAGGACAGAGAGACTACCTGC 4471  
4116 GTCGGAGGCGCAAGCTCTTAAAGCGGAGGGGACGCTTCGCCCCACCGCTCTACCGG 4175  
4470 GTCGGAGGCGCAAGCTCTTAAAGCGGAGGGGACGCTTCGCCCCACCGCTCTACCGG 4411  
4116 ACCTGACCGAGGCGCTCAAGAGGAGGAGCGCTGGGGCCCCGTGAAGGTGAAGCGGGCGCATG 4235  
4410 ACCTGACCGAGGCGCTCAAGAGGAGGAGCGCTGGGGCCCCGTGAAGGTGAAGCGGGCGCATG 4351  
4236 AGGGCTGTGTGGCCACGGTGAAGGAGGGGGCGGCTCATTCATGAGAGTCCGGCGGAGG 4295  
4350 AGGGCTGTGTGGCCACGGTGAAGGAGGGGGCGGCTCATTCATGAGAGTCCGGCGGAGG 4291  
4296 AGCTGCGGACACGCTCCGAGCTGCCCCCTGGCCCCCGGCTCAAGAGGGCTCATCA 4355  
4290 AGCTGCGGACACGCTCCGAGCTGCCCCCTGGCCCCCGGCTCAAGAGGGCTCATCA 4231  
4356 GCGAGGGGACCCCGCTCAAGTACGACACCGGGGGCTCCACCATGGCTCAAAAAGCACG 4415  
4230 GCGAGGGGACCCCGCTCAAGTACGACACCGGGGGCTCCACCATGGCTCAAAAAGCACG 4171  
4416 AGCTACGCTCCCTCATCGGACAGCCCCCGCGGACGTTCCACCCGCTGACCCGCTGATG 4475  
4170 AGCTACGCTCCCTCATCGGACAGCCCCCGCGGACGTTCCACCCGCTGACCCGCTGATG 4111  
4476 TGATGGCCGAGCGCCCGGGGACTGTGAACGTGCTGCTACGAGGAGAGCTGAAGCGCGC 4535  
4110 TGATGGCCGAGCGCCCGGGGACTGTGAACGTGCTGCTACGAGGAGAGCTGAAGCGCGC 4051  
4556 CAGGGACCGCCGACGCTCGGGGGGCTCATTTGCCGCGGGGCCCCGGTCATTGGGCTGG 4595  
4050 CAGGGACCGCCGACGCTCGGGGGGCTCATTTGCCGCGGGGCCCCGGTCATTGGGCTGG 3991  
4596 AGCTGGGTAAAGCCGCGGAGAGAGCCCCCTGACCTTATGAGGACCAAGGGGACCTTTGGCG 4655  
3990 AGCTGGGTAAAGCCGCGGAGAGAGCCCCCTGACCTTATGAGGACCAAGGGGACCTTTGGCG 3931  
4656 GCCACCTCCACAGAGGTTGCGCCGTGACCATGCGGAGGCCACGCGCGGCTGACAGAGG 4715  
3930 GCCACCTCCACAGAGTTCGCCCCGTGACCATGCGGAGGCCACGCGCGGCTGACAGAGG 3871  
4716 GCAGCTTTGCTGACAGCAAGGATCCAGGAGCGAAAGCTGAGGTGAGCGCTGCTGAGGA 4775  
3870 GCAGCTTTGCTGACAGCAAGGATCCAGGAGCGAAAGCTGAGGTGAGCGCTGCTGAGGA 3811  
4776 TCGCCAGTCTCCCGGACAGCACGTGCGCGGAGCACCAACCCCACTTCGCTGACTATG 4835  
3810 TCGCCAGTCTCCCGGACAGCACGTGCGCGGAGCACCAACCCCACTTCGCTGACTATG 3751  
4836 AGCACCCTGCTTGGGGCGTGAAGTGGCGTGAACCTGTATCGCAGCCACATCCCCCTGGCCT 4895  
|||||

Db 3750 AGCACCCTGCTTGGGGCGTGAAGTGGCGTGAACCTGTATCGCAGGCCACATCCCCCTGGCCT 3691  
Qy 4896 TCGACCCCACTCATACCCCGCGCATCCCTTGGAGCGAGCGCGCTCTACTACCTGC 4955  
Db 3690 TCGACCCCACTCATACCCCGCGCATCCCTTGGAGCGAGCGCGCTCTACTACCTGC 3631  
Qy 4956 CCGGACCTGGCCCCCAACCCCACTACCCGACCTGTACCCACCTTACCTCATCCGCG 5015  
Db 3630 CCGGACCTGGCCCCCAACCCCACTACCCGACCTGTACCCACCTTACCTCATCCGCG 3571  
Qy 5016 GCTACCCCGACACGGCGCGGCTGGAAGACCGGACAGCATCATGTAATGATCATCACT 5075  
Db 3570 GCTACCCCGACACGGCGCGGCTGGAAGACCGGACAGCATCATGTAATGATCATCACT 3511  
Qy 5076 CGCAGCAGATGACCAACAACGAGCCACCGCATGGCCAGCAGACTGATATGTGAAGG 5135  
Db 3510 CGCAGCAGATGACCAACAACGAGCCACCGCATGGCCAGCAGACTGATATGTGAAGG 3451  
Qy 5136 GCTTGTGCCCCGCGAGTCTGTGCTGSCATCTCAACTACGTGCGGGTCCCGAGGACTCA 5195  
Db 3450 GCTTGTGCCCCGCGAGTCTGTGCTGSCATCTCAACTACGTGCGGGTCCCGAGGACTCA 3391  
Qy 5196 TCGACCTGTCCCAAGTGGCAACACTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5255  
Db 3390 TCGACCTGTCCCAAGTGGCAACACTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3331  
Qy 5256 CCACCGCATGGAACGGCGCTGTACTCTCCCAACGCGGCCCAAGCGCTTACAGACCGCC 5315  
Db 3330 CCACCGCATGGAACGGCGCTGTACTCTCCCAACGCGGCCCAAGCGCTTACAGACCGCC 3271  
Qy 5316 ACAAGAGTCTCCCACTCTCCCAAGAGGTCCACAACATTGACAAACCAACCAACACGT 5375  
Db 3270 ACAAGAGTCTCCCACTCTCCCAAGAGGTCCACAACATTGACAAACCAACCAACACGT 3211  
Qy 5376 CCTGTCCGAGCGGAGGAGGAGACCGGGATCGAGAGCGGGACCGGGATGCGGGAGGAGAA 5435  
Db 3210 CCTGTCCGAGCGGAGGAGGAGACCGGGATCGAGAGCGGGACCGGGATGCGGGAGGAGAA 3151  
Qy 5436 AGTCCATCTCTACGTCACACGACGAGTGAAGCAGCACCCATCTGTGAGACTGTGTACAG 5495  
Db 3150 AGTCCATCTCTACGTCACACGAGCGTGAAGCAGCACCCATCTGTGAGACTGTGTACAG 3091  
Qy 5496 AGCAGAGACGGGACGACAGCGGACAGCGCGGGGGTGGGGGACAGACAGCGCGCCCG 5555  
Db 3090 AGCAGAGACGGGACGACAGCGGACAGCGCGGGGGTGGGGGACAGACAGCGCGCCCG 3031  
Qy 5556 CCTCCACTCCCATGCCCCAACAAGGACTGGCCCATCTCCCTCGGACCAAGATGCGCTCC 5615  
Db 3030 CCTCCACTCCCATGCCCCAACAAGGACTGGCCCATCTCCCTCGGACCAAGATGCGCTCC 2971  
Qy 5616 AGCAGAGACCGAGTGTCTTCAACAACAGGACATGAAGGATATCATCACCGCTGTGAGC 5675  
Db 2970 AGCAGAGACCGAGTGTCTTCAACAACAGGACATGAAGGATATCATCACCGCTGTGAGC 2911  
Qy 5676 CCAGCAAGCCCAAGGCTGTGAGGTCACTTCACCTCTCAACCGTGTGCCAGCTGCCA 5735  
Db 2910 CCAGCAAGCCCAAGGCTGTGAGGTCACTTCACCTCTCAACCGTGTGCCAGCTGCCA 2851  
Qy 5736 CATTTCCACTGTGCCACCACTGCCCCACTGGGCGGAGACCTGTGATGGGTCTAACCTTACC 5795  
Db 2850 CATTTCCACTGTGCCACCACTGCCCCACTGGGCGGAGACCTGTGATGGGTCTTACCTTACC 2791  
Qy 5796 TCATGAGACCGGTCTTGTGCTGCCAAGAGAGCCCCCGGGGTGCGCCGCGCAGAGCGGCCCC 5855  
Db 2790 TCATGAGACCGGTCTTGTGCTGCCAAGAGAGCCCCCGGGGTGCGCCGCGCAGAGCGGCCCC 2731  
Qy 5856 GAGCAGACACCGGCATGCTTCTGTGCAAGCCCCAGCCCCGCTTCGAGCTGAGCCCCG 5915  
Db 2730 GAGCAGACACCGGCATGCTTCTGTGCAAGCCCCAGCCCCGCTTCGAGCTGAGCCCCG 2671  
Qy 5916 CCTCTCTCCCCCAGAGGAGCTGAGAGCCCCCGGCTTGTGTGTCTGTCTTGTGCGACG 5975  
Db 2670 CCTCTCTCCCCCAGAGGAGCTGAGAGCCCCCGGCTTGTGTGTCTGTCTTGTGCGACG 2611  
|||||

QY 5976 CCAACATCGCCCGCACCCCTGCGAAGAACTCGCACTCAACAGCCGCGGACCCGCG 6035  
| | | | |  
Db 2610 CCAACATCGCCCGCACCCCTGCGAAGAACTCGCACTCAACAGCCGCGGACCCGCG 2551  
| | | | |  
QY 6036 CCGCGCCACCTGCTGCGCTCGGACCCGCGACCGGGAAAAAGACTCAAAAGTAAACCTTTT 6095  
| | | | |  
Db 2550 CCGCGCCACCTGCTGCGCTCGGACCCGCGACCGGGAAAAAGACTCAAAAGTAAACCTTTT 2491  
| | | | |  
QY 6096 CCAATCCAGGAACCTGGAACCTCCGTTCTGTGGTTAACCAACGCGACAGCTAACAGCCCGGAAG 6155  
| | | | |  
Db 2490 CCAATCCAGGAACCTGGAACCTCCGTTCTGTGGTTAACCAACGCGACAGCTAACAGCCCGGAAG 2431  
| | | | |  
QY 6156 GGGTGGAGCCCGCTCAGCCCTGTGAGCTCACCCAGTGAACCCAGCAAGGGGCTCCCA 6215  
| | | | |  
Db 2430 GGGTGGAGCCCGCTCAGCCCTGTGAGCTCACCCAGTGAACCCAGCAAGGGGCTCCCA 2371  
| | | | |  
QY 6216 AGCACTGGAAGAGCTGCAACAAGAGCCACTGGAAGGGGAGAGCTGCGGCCCAAGCAGCCAG 6275  
| | | | |  
Db 2370 AGCACTGGAAGAGCTGCAACAAGAGCCACTGGAAGGGGAGAGCTGCGGCCCAAGCAGCCAG 2311  
| | | | |  
QY 6276 GGGCCGTTGAAGCTTGGCGGGGGAGAGCCGCCCACTCCCAACCTGCGGGCGCTGCTGAGA 6335  
| | | | |  
Db 2310 GGGCCGTTGAAGCTTGGCGGGGGAGAGCCGCCCACTCCCAACCTGCGGGCGCTGCTGAGA 2251  
| | | | |  
QY 6336 GCGAGCCCTGCTCAGCCCGCTGCTCAGAGCCGCCCGAGGGGTCAAAAGTCAACAGCGGG 6395  
| | | | |  
Db 2250 GCGAGCCCTGCTCAGCCCGCTGCTCAGAGCCGCCCGAGGGGTCAAAAGTCAACAGCGGG 2191  
| | | | |  
QY 6396 TGGTCACTCTGAGCCAGCAATCACTGAGAGTCACTCAACAGAGCTAACCCGCGACACC 6455  
| | | | |  
Db 2190 TGGTCACTCTGAGCCAGCAATCACTGAGAGTCACTCAACAGAGCTAACCCGCGACACC 2131  
| | | | |  
QY 6456 CACAGAGACTCAGCCGCAACCCCTGCGCCGCCCTCTACTCTCTTCCCTGGGGCCAGCTGCC 6515  
| | | | |  
Db 2130 CACAGAGACTCAGCCGCAACCCCTGCGCCGCCCTCTACTCTCTTCCCTGGGGCCAGCTGCC 2071  
| | | | |  
QY 6516 CGTCTGAGACCTCGCGCGGCCCAACAGTGAACCTTAACCTCCCGGCCCGGACCATGGTG 6575  
| | | | |  
Db 2070 CGTCTGAGACCTCGCGCGGCCCAACAGTGAACCTTAACCTCCCGGCCCGGACCATGGTG 2011  
| | | | |  
QY 6576 CCGCGGCGCGGTGGCTCCCGCAACGGAAGGGGCAAGAGTCTCCAGAGCCCAACAGAG 6635  
| | | | |  
Db 2010 CCGCGGCGCGGTGGCTCCCGCAACGGAAGGGGCAAGAGTCTCCAGAGCCCAACAGAG 1951  
| | | | |  
QY 6636 CGTCCGTCTTGGGGTGGTGGTGAAGAGGTTGAACCTGTGTCCCAACGAGGGGCAATGA 6695  
| | | | |  
Db 1950 CGTCCGTCTTGGGGTGGTGGTGAAGAGGTTGAACCTGTGTCCCAACGAGGGGCAATGA 1891  
| | | | |  
QY 6696 CCGAGCCAGGGGACCTCCGGAAGTGTGTGAACCCGCTGCTGTACCGGGAGTGGGGACAGA 6755  
| | | | |  
Db 1890 CCGAGCCAGGGGACCTCCGGAAGTGTGTGAACCCGCTGCTGTACCGGGAGTGGGGACAGA 1831  
| | | | |  
QY 6756 CCGAGCCAGAGAGTGGGCTCAAGTCTTCAGGCAACACAGCCAGCCGCGGACCTTCT 6815  
| | | | |  
Db 1830 CCGAGCCAGAGAGTGGGCTCAAGTCTTCAGGCAACACAGCCAGCCGCGGACCTTCT 1771  
| | | | |  
QY 6816 TTAGCAAGCTGACCGAGAGCAACTCCGCAATGGTCAAGTCCAGAGAGCAAGAGATCAAC 6875  
| | | | |  
Db 1770 TTAGCAAGCTGACCGAGAGCAACTCCGCAATGGTCAAGTCCAGAGAGCAAGAGATCAAC 1711  
| | | | |  
QY 6876 AGAAGCTGAACCCCAACACCGGAATGAGCTGAATACAAATATACAGCCAGCCCTGGGACGG 6935  
| | | | |  
Db 1710 AGAAGCTGAACCCCAACACCGGAATGAGCTGAATACAAATATACAGCCAGCCCTGGGACGG 1651  
| | | | |  
QY 6936 AGATCTTCAATATGCGCGCATCAACCGGAACAGGCTTATGACTTATAGAGCCAGCGCG 6995  
| | | | |  
Db 1650 AGATCTTCAATATGCGCGCATCAACCGGAACAGGCTTATGACTTATAGAGCCAGCGCG 1591  
| | | | |  
QY 6996 TGCAGGAACATGCGAGACCAACATGGGGCTGAGGCCATTAATGAAAGGCACTCATGG 7055  
| | | | |  
Db 1590 TGCAGGAACATGCGAGACCAACATGGGGCTGAGGCCATTAATGAAAGGCACTCATGG 1531  
| | | | |

QY 7056 GTAATATATGACAGTGGGAAGAGTCCCGCGCTCAGCGCCAAATGCTTTTAACCTCTGA 7115  
| | | | |  
Db 1530 GTAATATATGACAGTGGGAAGAGTCCCGCGCTCAGCGCCAAATGCTTTTAACCTCTGA 1471  
| | | | |  
QY 7116 ATGCGAGTGCAGGCTGCGCGCTGCTATGCTCAATTAACCGTGTGACGAGCGAGTGAAC 7175  
| | | | |  
Db 1470 ATGCGAGTGCAGGCTGCGCGCTGCTATGCTCAATTAACCGTGTGACGAGCGAGTGAAC 1411  
| | | | |  
QY 7176 ACACACTCACTTCGCAAGGTGGCGGGGAAAGGCCAAAGGCTCTGGCAGACCCAGCAGCC 7235  
| | | | |  
Db 1410 ACACACTCACTTCGCAAGGTGGCGGGGAAAGGCTCTGGCAGACCCAGCAGCC 1351  
| | | | |  
QY 7236 GAAAGCCAAATGCTCCCGGCCCGCGGCTGAGCTTGGAGAACCGGCCACCTCTGTCTCT 7295  
| | | | |  
Db 1350 GAAAGCCAAATGCTCCCGGCCCGCGGCTGAGCTTGGAGAACCGGCCACCTCTGTCTCT 1291  
| | | | |  
QY 7296 CAGTGACTTCGAGAGGAGACTGCAACCGCCGAGCGCTCAGCAACCCGCTGTGGAGG 7355  
| | | | |  
Db 1290 CAGTGACTTCGAGAGGAGACTGCAACCGCCGAGCGCTCAGCAACCCGCTGTGGAGG 1231  
| | | | |  
QY 7356 ACAGGCGCTCGTCCGAGGTTCCAGCGCCATCCCTTACAAACCCCGTATCATGGGCTGC 7415  
| | | | |  
Db 1230 ACAGGCGCTCGTCCGAGGTTCCAGCGCCATCCCTTACAAACCCCGTATCATGGGCTGC 1171  
| | | | |  
QY 7416 AGCGGGGTGTCATGAGCTTCCCAACCCCAACCGGAGCTCCCGCGGCGAGCGGCGCTCG 7475  
| | | | |  
Db 1170 AGCGGGGTGTCATGAGCTTCCCAACCCCAACCGGAGCTCCCGCGGCGAGCGGCGCTCG 1111  
| | | | |  
QY 7476 CTGGCCCCCAACAACGCTTGGAGAGAGAGCCCAAGCCACTGCTCTGTCTGAGTACAGAG 7535  
| | | | |  
Db 1110 CTGGCCCCCAACAACGCTTGGAGAGAGAGCCCAAGCCACTGCTCTGTCTGAGTACAGAG 1051  
| | | | |  
QY 7536 CACTCTCCGACACCGAGTGACTCAAGAACAGGCGGGGGGGGGGGCGGCTGCTGCTGCTG 7595  
| | | | |  
Db 1050 CACTCTCCGACACCGAGTGACTCAAGAACAGGCGGGGGGGGGGGCGGCTGCTGCTGCTG 1001  
| | | | |  
QY 7596 GTCAAGGTCCCAAGAGCCACAAGAAACGAGCCCTGACAGAGACGGGGCGAGTGCAGACTCCC 7655  
| | | | |  
Db 1000 GTCAAGGTCCCAAGAGCCACAAGAAACGAGCCCTGACAGAGACGGGGCGAGTGCAGACTCCC 941  
| | | | |  
QY 7656 CAACCAAGGAAGAGCCCTGAGTCCGCTGCGCTCCATCCATCTGTCCGTCCAGAGCC 7715  
| | | | |  
Db 940 CAACCAAGGAAGAGCCCTGAGTCCGCTGCGCTCCATCCATCTGTCCGTCCAGAGCC 881  
| | | | |  
QY 7716 GGCATCTCTTCCCTGCTTAAAGCTTAACTTAAGCTCCCGCCCGGCTGGCTGGCGAG 7775  
| | | | |  
Db 880 GGCATCTCTTCCCTGCTTAAAGCTTAACTTAAGCTCCCGCCCGGCTGGCTGGCGAG 821  
| | | | |  
QY 7776 ACCTTACTCAGGGAGTGTTAACCTGGTGTCCGGAAGGGAAGGGGACCGGGAGGG 7835  
| | | | |  
Db 820 ACCTTACTCAGGGAGTGTTAACCTGGTGTCCGGAAGGGAAGGGGACCGGGAGGG 761  
| | | | |  
QY 7836 GGCACGCGAGCGGTGTGGAGCCACAACAAGCGGCGCAGGGCGGCCAGAGGACCCAAAGCA 7895  
| | | | |  
Db 760 GGCACGCGAGCGGTGTGGAGCCACAACAAGCGGCGCAGGGCGGCCAGAGGACCCAAAGCA 701  
| | | | |  
QY 7896 GGAATGACCAAGCACTTCCACGCGCACTGCTCCCGGAAATGCAATTTTGAACCAAAAGCTTAA 7955  
| | | | |  
Db 700 GGAATGACCAAGCACTTCCACGCGCACTGCTCCCGGAAATGCAATTTTGAACCAAAAGCTTAA 641  
| | | | |  
QY 7956 ACTGAGCTGCAAGCCCGCGGCTCCCTCCGCTCCGCTCCCATCCCGCTTAAAGCTTGAACA 8015  
| | | | |  
Db 640 ACTGAGCTGCAAGCCCGCGGCTCCCTCCGCTCCGCTCCCATCCCGCTTAAAGCTTGAACA 581  
| | | | |  
QY 8016 GATGAGCGAGGCGCTGTGCAGGCCCCCAAGTGCCTGTTCGAGTCCCAACAAGCTGCCCC 8075  
| | | | |  
Db 580 GATGAGCGAGGCGCTGTGCAGGCCCCCAAGTGCCTGTTCGAGTCCCAACAAGCTGCCCC 521  
| | | | |  
QY 8076 AGCCAAACGAATTTGCTGGAACCAAGTCAAGGCGCAAGTGGCGGACAAAGGGCGCAGGTGC 8135  
| | | | |  
Db 520 AGCCAAACGAATTTGCTGGAACCAAGTCAAGGCGCAAGTGGCGGACAAAGGGCGCAGGTGC 461  
| | | | |  
QY 8136 GGCCTGGGGGGAACGGAATGCTCCGAGGACTGGAAGTGTTTTTCACACATCGTTGCCGCA 8195  
| | | | |

Db	460	GGCTTGGGGGAAACGATGCTCCAGAGACTGACTGTTTTTTTACACATGTCGCGCA	401
Qy	8196	GCGGTGGAAAGAAAAGCAGATGTAAATGATGTGTGTTTACAGGGTATATTTTGATA	8255
Db	400	GCGGTGGGAAAGAAAGCAGATGTAAATGATGTGTGTTTACAGGGTATATTTTGATA	341
Qy	8256	CCCTCAATGAATTAATTCAGATGTTTTACGCAAGGAGACTTACCAAGTATCTGCTG	8315
Db	340	CCCTCAATGAATTAATTCAGATGTTTTACGCAAGGAGACTTACCCAGTATCTGCTG	281
Qy	8316	CTGTGCTTTTGATCTCTGCTTACCGTTCAAGAGCGTGTGAGGCGCAACTGCGTACC	8375
Db	280	CTGTGCTTTTGATCTCTGCTTACCGTTCAAGAGCGTGTGAGGCGCAACTGCGTACC	221
Qy	8376	CCATCACTCGAGAGACCAAGGGGGGGGGAGCTGCTCG-TCAAGGCCCGCTGTGCTCCC	8434
Db	220	CCATCACTCGAGAGACCAAGGGGGGGGGAGCTGCTCGCTCACGCCCGCTGTGCTCCC	161
Qy	8435	TCCCTCCCTTCTTGGGAGAGTAAATTCGATGCGTATCTGTGGCGCCATTGGCGAG	8494
Db	160	TCCCTCCCTTCTTGGGAGAGTAAATTCGATGCGTATCTGTGGCGCCATTGGCGAG	101
Qy	8495	GGTGTGTGATTTCTGTCAATTACACACGTCGTTCTTAATTAAGGCAATTATCTCCAA	8554
Db	100	GGTGTGTGATTTCTGTCAATTACACACGTCGTTCTTAATTAAGGCAATTATCTCCAA	41
Qy	8555	AAAAAAAAAAAAAAAAA 8571	
Db	40	AAAAAAAAAAAAAAAAA 24	

CC	XX	RESULT 13
CC	XX	ADJ92815
CC	ID	ADJ92815 standard; DNA; 7554 BP.
CC	AC	ADJ92815;
CC	DT	06-MAY-2004 (first entry)
CC	DE	Human co-repressor SMRT DNA.
CC	KW	Nuclear receptor; therapy; diabetes; hyperlipidaemia; obesity;
CC	XX	atherosclerosis; human; co-repressor; gene; ds.
CC	OS	Homo sapiens.
CC	PN	US200328607-A1.
CC	PD	11-DEC-2003.
CC	PP	14-APR-2003; 2003US-00414692.
CC	PR	15-APR-2002; 2002US-0372650P.
CC	PA	(WAGN/) WAGNER B L.
CC	PA	(SCHU/) SCHULMAN I G.
CC	PT	Wagner BL, Schulman IG;
CC	DR	WPI; 2004-167207/16.
CC	PT	Identifying compounds that bind to nuclear receptor and exhibit cell type
CC	PT	specific actions, and useful for treating hyperlipidemia, obesity and
CC	XX	diabetes.
CC	PS	Claim 5; SEQ ID NO 14; 99pp; English.
CC	XX	The invention relates to screening methods for identifying compounds that
CC	XX	bind to nuclear receptor and exhibit cell type specific actions. The
CC	XX	invention relates to modulators having an improved therapeutic profile.
CC	XX	The method is useful for identifying compounds that bind to a nuclear
CC	XX	receptor and exhibit cell type specific actions. It is also useful for

CC identifying modulators of nuclear receptors that are useful in treating  
CC diseases e.g., diabetes, hyperlipidaemia, obesity, atherosclerosis, etc.  
CC The present sequence is human co-repressor DNA used to illustrate the  
CC method of the invention.

XX

SQ Sequence 7554 BP; 1646 A; 2732 C; 2228 G; 948 T; 0 U; 0 Other;

Query Match	88.2%	Score 7554;	DB 12;	Length 7554;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 7554;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	2	ATGTGCGGGCTCAACAAGCTGTGGACACAGCGTGGAAAGGGCACCTGAACCCCGCTAACCG	61
Db	1	ATGTGCGGGCTCAACAAGCTGTGGACACAGCGTGGAAAGGGCACCTGAACCCCGCTAACCG	60
QY	62	CCCCACAGCCCTTCTCTACCAAGTCAGATCGCCCGGACGCAACGGACGTGCGGCTCTG	121
Db	61	CCCCACAGCCCTTCTCTACCAAGTCAGATCGCCCGGACGCAACGGACGTGCGGCTCTG	120
QY	122	GAGTACCAAGACCACTCCCGGACTATAGCTCTCCACCTGTGTGGCGGGCTCCATCATCCAG	181
Db	121	GAGTACCAAGACCACTCCCGGACTATAGCTCTCCACCTGTGTGGCGGGCTCCATCATCCAG	180
QY	182	CCCCAGCGGGCGGAAGGCGCTCCCTGCTGCTGTAGATTCCAGCCCGGGAATGAACGTCGCCAG	241
Db	181	CCCCAGCGGGCGGAAGGCGCTCCCTGCTGCTGTAGATTCCAGCCCGGGAATGAACGTCGCCAG	240
QY	242	GAGCTCCACTGTGCGGCGCAGAGTCCCACTCATACCTGCCGAGCTGGGGAATCAGAGTGC	301
Db	241	GAGCTCCACTGTGCGGCGCAGAGTCCCACTCATACCTGCCGAGCTGGGGAATCAGAGTGC	300
QY	302	GAGTTCAATTGAACCAAGCGGCGCTCGGTAAAGCGTGCCTGACCCCTGCTGGACCG	361
Db	301	GAGTTCAATTGAACCAAGCGGCGCTCGGTAAAGCGTGCCTGACCCCTGCTGGACCG	360
QY	362	TCACCCCTGTGCGGCGCACGCGGCGCAGCTGTGCGGGAATCTGAAGACTTCACCAAGGACCGTAGC	421
Db	361	TCACCCCTGTGCGGCGCACGCGGCGCAGCTGTGCGGGAATCTGAAGACTTCACCAAGGACCGTAGC	420
QY	422	CTGAACGCGGAAGCTGGAAACCGGTGTCTTCCCCCAGCCCCCGGCACACTGACCTGTAGCTG	481
Db	421	CTGAACGCGGAAGCTGGAAACCGGTGTCTTCCCCCAGCCCCCGGCACACTGACCTGTAGCTG	480
QY	482	GAGCTGGTGGCGGCACCGGCTGTCCAAAGGAGAGCTGATCCAAACATGAGCCGCGTGGAC	541
Db	481	GAGCTGGTGGCGGCACCGGCTGTCCAAAGGAGAGCTGATCCAAACATGAGCCGCGTGGAC	540
QY	542	CGAGAGATCAACCATGTGTAGACAGCAGATCTCTAAGCTGAAGAAAGACAGCAACAGCTG	601
Db	541	CGAGAGATCAACCATGTGTAGACAGCAGATCTCTAAGCTGAAGAAAGACAGCAACAGCTG	600
QY	602	GAGAGAGAGGCTGCCAAGCCGCGCCGAGGCTTGAGAAAGCCCGTGTACACCGCCGCCATTCGAG	661
Db	601	GAGAGAGAGGCTGCCAAGCCGCGCCGAGGCTTGAGAAAGCCCGTGTACACCGCCGCCATTCGAG	660
QY	662	TCGAAGCACCGCAGGCTGTGTGCAATCATCTACGACGGAACCGGAAAGAGGCTAAGCT	721
Db	661	TCGAAGCACCGCAGGCTGTGTGCAATCATCTACGACGGAACCGGAAAGAGGCTAAGCT	720
QY	722	GCACATCGGAATTCGGAAGGCTGTGGGCGCCAGGTGAGCTGCGCTGTACAAACGAGCC	781
Db	721	GCACATCGGAATTCGGAAGGCTGTGGGCGCCAGGTGAGCTGCGCTGTACAAACGAGCC	780
QY	782	TCCGACACCCCGCAGTATCATGAGAAATCAAAATTAACCAAGCGGATCCGGAAGAAAGCTA	841
Db	781	TCCGACACCCCGCAGTATCATGAGAAATCAAAATTAACCAAGCGGATCCGGAAGAAAGCTA	840
QY	842	ATCTTGTACTTCAAGAGGAGGATTCACGCTCGGAAACATAGGAACGAAGTTCTGCGAG	901
Db	841	ATCTTGTACTTCAAGAGGAGGATTCACGCTCGGAAACATAGGAACGAAGTTCTGCGAG	900
QY	902	CGCTATGACACAGCTCATGAGGCTCTTGAAAAAAAAGTGGAGCGCATCGAAAAACAACCG	961

Dh 901 CGCTATGACCAAGCTCATGAGAGCGCTTGGAAAAAAGGTGAGCGCATCGAAAACAAACCG 960  
Qy 962 CGCGCGCGGGCCCAAGGAGCAAGGTGCGGAGTACTACGAAAAAGAGTTCCCTGAGATC 1021  
Db 961 CGCGCGCGGGCCCAAGGAGCAAGGTGCGGAGTACTACGAAAAAGAGTTCCCTGAGATC 1020  
Qy 1022 CGCAAGCAGCGGAGCTGCAAGAGCGCATATGCAAGAGAGGTGGCCAGCGGGCAAGTGG 1081  
Db 1021 CCAGAGCAGCGGAGCTGCAAGAGCGCATATGCAAGAGAGGTGGCCAGCGGGCAAGTGG 1080  
Qy 1082 CTGTCCATGTGGCGCGCGCGCAGCGAGCAAGAGGTGCAAGATCATTCGATGGCTCTCA 1141  
Db 1081 CTGTCCATGTGGCGCGCGCGCAGCGAGCAAGAGGTGCAAGATCATTCGATGGCTCTCA 1140  
Qy 1142 GAGCAGGAGAACTGTGAGAAAGAGATGCGCAGCTGGCGGTGATCCCGCCATGCTGTAC 1201  
Db 1141 GAGCAGGAGAACTGTGAGAAAGAGATGCGCAGCTGGCGGTGATCCCGCCATGCTGTAC 1200  
Qy 1202 GACGCTGACCAAGCAGCGCATCAAGTTCAATCAATGAAAGGGCTTATGGCCGACCCCATG 1261  
Db 1201 GACGCTGACCAAGCAGCGCATCAAGTTCAATCAATGAAAGGGCTTATGGCCGACCCCATG 1260  
Qy 1262 AAGGTGTACAAAGACCGCGAGGTCAATGAAATGTGAGTGAAGAGAGAGACCTTC 1321  
Db 1261 AAGGTGTACAAAGACCGCGAGGTCAATGAAATGTGAGTGAAGAGAGAGACCTTC 1320  
Qy 1322 CGGGAAGATTGATGAGATCCCAAGAACTTGGGCTGATGCGCATCATTCCTGGAGAG 1381  
Db 1321 CGGGAAGATTGATGAGATCCCAAGAACTTGGGCTGATGCGCATCATTCCTGGAGAG 1380  
Qy 1382 AAGACAGTGGTGAAGTGCCTCTCTATTACTACTGACTTAAGAAATGAGAACTAAG 1441  
Db 1381 AAGACAGTGGTGAAGTGCCTCTCTATTACTACTGACTTAAGAAATGAGAACTAAG 1440  
Qy 1442 AGCTTGTGAGACGAGACTATTCGGCGCGCGCAAGAGCCAGCAAGCAACAGCAGCAG 1501  
Db 1441 AGCTTGTGAGACGAGACTATTCGGCGCGCGCAAGAGCCAGCAAGCAACAGCAGCAG 1500  
Qy 1502 CAGCAGCAGCAAGCAG 1561  
Db 1501 CAGCAGCAGCAAGCAG 1560  
Qy 1562 GATGAG 1621  
Db 1561 GATGAG 1620  
Qy 1622 GACAAAGAAAGACTCTCTCAAGAGAAAGACAGACACTCAGGGAGAGCAACGACGAG 1681  
Db 1621 GACAAAGAAAGACTCTCTCAAGAGAAAGACAGACACTCAGGGAGAGCAACGACGAG 1680  
Qy 1682 AAGGAGGCTGTGGCTCTCCAAAGGCGGCAAACTGCCAACAGCCAGGGAAAGCGCAAGGC 1741  
Db 1681 AAGGAGGCTGTGGCTCTCCAAAGGCGGCAAACTGCCAACAGCCAGGGAAAGCGCAAGGC 1740  
Qy 1742 CGCATCAACCGGCTCAATGCTATATGAGGCGCAAGGAGAGAGGCGCATCCCGCCAGCAG 1801  
Db 1741 CGCATCAACCGGCTCAATGCTATATGAGGCGCAAGGAGAGAGGCGCATCCCGCCAGCAG 1800  
Qy 1802 AGCGCCGAGCTGGCTCTCAATGAGCTGAATGAAGTTCTCGCTGAGCAGAAAGAAATG 1861  
Db 1801 AGCGCCGAGCTGGCTCTCAATGAGCTGAATGAAGTTCTCGCTGAGCAGAAAGAAATG 1860  
Qy 1862 GAAACAGCCAAAGAAAGTCTCTTGAAGACGCGCGGCACTGGTCCGCTCCCGGAGT 1921  
Db 1861 GAAACAGCCAAAGAAAGTCTCTTGAAGACGCGCGGCACTGGTCCGCTCCCGGAGT 1920  
Qy 1922 GTGGGCTCCAAAGCTGTGCGAGTGAAGAACTTACTTCACTTCACTCAAGAAAGAGCAG 1981  
Db 1921 GTGGGCTCCAAAGCTGTGCGAGTGAAGAACTTACTTCACTTCACTCAAGAAAGAGCAG 1980  
Qy 1982 AACCTTGATGAGATCTTGCAGAGCAAGAGCTGAAGATGAGAGAGAGAGAGAGCGCGCG 2041  
Db 1981 AACCTTGATGAGATCTTGCAGAGCAAGAGCTGAAGATGAGAGAGAGAGAGAGCGCGCG 2040

Qy 2042 AGGAAAGAAAGAAAGCGCGCGCGCGCGCAGCGAGAGAGCTGCAATTCCCGCCGCTGTG 2101  
Db 2041 AGGAAAGAAAGAAAGCGCGCGCGCGCGCGCAGCGAGAGAGCTGCAATTCCCGCCGCTGTG 2100  
Qy 2102 GAGGATGAGAGATGAGAGGCTGTGGCGGTGACGCGAAATGAGAGAGAGATGTGAGAGAG 2161  
Db 2101 GAGGATGAGAGATGAGAGGCTGTGGCGGTGACGCGAAATGAGAGAGAGATGTGAGAGAG 2160  
Qy 2162 GCTGAAGCCTTACATGCGCTCTGAGAAATGAGGTGCCAGAGGGGAAATGCAATGGCCCAAGC 2221  
Db 2161 GCTGAAGCCTTACATGCGCTCTGAGAAATGAGGTGCCAGAGGGGAAATGCAATGGCCCAAGC 2220  
Qy 2222 ACTGTCAACAAAGCTCAGACACCGAGACATCCCTCTCTCACTGAGCGGCGCAAG 2281  
Db 2221 ACTGTCAACAAAGCTCAGACACCGAGACATCCCTCTCTCACTGAGCGGCGCAAG 2280  
Qy 2282 GACACAGGGCAGAAATGGGCTCAAGCCCGCAAGCCACTTGGGCGCGAGCGGGCAACCCCA 2341  
Db 2281 GACACAGGGCAGAAATGGGCTCAAGCCCGCAAGCCACTTGGGCGCGAGCGGGCAACCCCA 2340  
Qy 2342 GGCCTCACCCACCCACCAAGAGAGACATCCCGGGCCCCATTGAGGCCACCCGGGCTCT 2401  
Db 2341 GGCCTCACCCACCCACCAAGAGAGACATCCCGGGCCCCATTGAGGCCACCCGGGCTCT 2400  
Qy 2402 GAAGCCACCGAGCGCTTACGCGCCCCACAGCACCCCCCATGGCCCTCTGCACTCTCTCT 2461  
Db 2401 GAAGCCACCGAGCGCTTACGCGCCCCACAGCACCCCCCATGGCCCTCTGCACTCTCTCT 2460  
Qy 2462 GTGTGTCCCAAG 2521  
Db 2461 GTGTGTCCCAAG 2520  
Qy 2522 GAGAGAGAGAAAGCCCCCGCGCTGAGAGAGCTGTGACGTGACACAGAGGAAAGGCCAGAGAG 2581  
Db 2521 GAGAGAGAGAAAGCCCCCGCGCTGAGAGAGCTGTGACGTGACACAGAGGAAAGGCCAGAGAG 2580  
Qy 2582 CCGGTCAAGAGAGAGTGCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2641  
Db 2581 CCGGTCAAGAGAGAGTGCACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
Qy 2642 GAGGCGCTGAGAGGCAACGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2701  
Db 2641 GAGGCGCTGAGAGGCAACGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700  
Qy 2702 GGCAGGAGCCACACTGACCAAGAGCTGTGGGCGCGCCCCAGAGACAGGCACTTCAGTGTACC 2761  
Db 2701 GGCAGGAGCCACACTGACCAAGAGCTGTGGGCGCGCCCCAGAGACAGGCACTTCAGTGTACC 2760  
Qy 2762 TGCAGTGCAGAGAGGTGAGTGAAGGCGAGGGCGAGCAAGAAACGGGCTGTGTCCCA 2821  
Db 2761 TGCAGTGCAGAGAGGTGAGTGAAGGCGAGGGCGAGCAAGAAACGGGCTGTGTCCCA 2820  
Qy 2822 AGGCGCAGGCTCTCTCAACCCGACTGGCGAGCCCCGGGCGCAATGCTCAACCCAGAAAGCA 2881  
Db 2821 AGGCGCAGGCTCTCTCAACCCGACTGGCGAGCCCCGGGCGCAATGCTCAACCCAGAAAGCA 2880  
Qy 2882 CTGAGACTGAGCAAGCTGAAGCAGGAGGCTGTGCCATTCCTCCCATTCAGAGTCAACAAA 2941  
Db 2881 CTGAGACTGAGCAAGCTGAAGCAGGAGGCTGTGCCATTCCTCCCATTCAGAGTCAACAAA 2940  
Qy 2942 GTTCATGAGCGCGCGCGGAGAGACGAGCTCCACCAAGCAGGCTCCCGAGCCCAAGC 3001  
Db 2941 GTTCATGAGCGCGCGCGGAGAGACGAGCTCCACCAAGCAGGCTCCCGAGCCCAAGC 3000  
Qy 3002 CCAACGCAAAACCTTGAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3061  
Db 3001 CCAACGCAAAACCTTGAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
Qy 3062 GGCAG 3121  
Db 3061 GGCAG 3120

QY 3122 AAGTGCTGGAGACCCCTTGCTGAATTCCGGCTGCTCCCTTCCCGTCCCCCGT 3181  
| | | | |  
Db 3121 AAGTGCTGGAGACCCCTTGCTGAATTCCGGCTGCTCCCTTCCCGTCCCCCGT 3180  
| | | | |  
QY 3182 GAGGATCAAGAGCCTCCCGCATGCCCGGACCCCTGAGCCTTCTCAAGCTCAACT 3241  
| | | | |  
Db 3181 GAGGATCAAGAGCCTCCCGCATGCCCGGACCCCTGAGCCTTCTCTCAAGCTCAACT 3240  
| | | | |  
QY 3242 GGTCAACCACTGCCCCCTGGGCTTCATGACATGCCCCGGCCGTCGCGCGCCACCC 3301  
| | | | |  
Db 3241 GGTCAACCACTGCCCCCTGGGCTTCATGACATGCCCCGGCCGTCGCGCGCCACCC 3300  
| | | | |  
QY 3302 ACCATCTCAACCCGCTCTCCCTCATTTCTCTGCGCAAGCACCCGAGCTCTCAAGG 3361  
| | | | |  
Db 3301 ACCATCTCAACCCGCTCTCCCTCATTTCTCTGCGCAAGCACCCGAGCTCTCAAGG 3360  
| | | | |  
QY 3362 CAATAAGGTGCATCTCCCAAGGAATGTCCGTCCAGCTCCAGTCCCGTACTCAAGCAT 3421  
| | | | |  
Db 3361 CAATAAGGTGCATCTCCCAAGGAATGTCCGTCCAGCTCCAGTCCCGTACTCAAGCAT 3420  
| | | | |  
QY 3422 GCCAAGCCCCGGTGGGCTCTGTCAACATGGGGCTGCCCTGCCATGGAACCCAAAAAG 3481  
| | | | |  
Db 3421 GCCAAGCCCCGGTGGGCTCTGTCAACATGGGGCTGCCCTGCCATGGAACCCAAAAAG 3480  
| | | | |  
QY 3482 CTGSCACCTTTAGCTGAGTGAAGAGAGAGAGCTGTCTCCACGCGGGCCAGGCTGGGCCA 3541  
| | | | |  
Db 3481 CTGSCACCTTTAGCTGAGTGAAGAGAGAGAGCTGTCTCCACGCGGGCCAGGCTGGGCCA 3540  
| | | | |  
QY 3542 CCGGAGAGCTGGGGGGTCCCAAGCCCAAGAGGGGTCCGTGTGAGAGAGAGCACTCTCG 3601  
| | | | |  
Db 3541 CCGGAGAGCTGGGGGGTCCCAAGCCCAAGAGGGGTCCGTGTGAGAGAGAGCACTCTCG 3600  
| | | | |  
QY 3602 GGCTCAGTTCCGGGCGGAAGCATCAACAAAGGCATTTCCAGACACAGGGTGCCTCGGAC 3661  
| | | | |  
Db 3601 GGCTCAGTTCCGGGCGGAAGCATCAACAAAGGCATTTCCAGACACAGGGTGCCTCGGAC 3660  
| | | | |  
QY 3662 AGCGCATCACTACCGCGGGCTTCATCAACCGGACCGGACGTCGCTGTACAAG 3721  
| | | | |  
Db 3661 AGCGCATCACTACCGCGGGCTTCATCAACCGGACCGGACGTCGCTGTACAAG 3720  
| | | | |  
QY 3722 GGCAATCAACCAAGATCATTCGGCGAGGACAGCCGAGTCCGTGTGAGACCGCGGGCGAG 3781  
| | | | |  
Db 3721 GGCAATCAACCAAGATCATTCGGCGAGGACAGCCGAGTCCGTGTGAGACCGCGGGCGAG 3780  
| | | | |  
QY 3782 GACAGCTTGCCCAAGGGCCACGTCACTCAAGAGGCAAGAGGGCCACGTCCTTGCTCAT 3841  
| | | | |  
Db 3781 GACAGCTTGCCCAAGGGCCACGTCACTCAAGAGGCAAGAGGGCCACGTCCTTGCTCAT 3840  
| | | | |  
QY 3842 GAGGATGATGTCTGTGAACCAAGTCTCAAGAGAGAGCGGACAGAGCATCAGAGACC 3901  
| | | | |  
Db 3841 GAGGATGATGTCTGTGAACCAAGTCTCAAGAGAGAGCGGACAGAGCATCAGAGACC 3900  
| | | | |  
QY 3902 CCCCATGAGAGCGCGCCCCCAAGGCAACCTATGACATGATGAGAGGCGCGCTGGGACAG 3961  
| | | | |  
Db 3901 CCCCATGAGAGCGCGCCCCCAAGGCAACCTATGACATGATGAGAGGCGCGCTGGGACAG 3960  
| | | | |  
QY 3962 GGCATCTCTGAGCAGCATCGAAGTCTCATGGGCGGTGCCATCCCGCGGAGCGACAC 4021  
| | | | |  
Db 3961 GGCATCTCTGAGCAGCATCGAAGTCTCATGGGCGGTGCCATCCCGCGGAGCGACAC 4020  
| | | | |  
QY 4022 AGCCCCCAACCTCAAGAGAGAGACCAATCCGCGGGTTCATCAACAAAGGATTCCT 4081  
| | | | |  
Db 4021 AGCCCCCAACCTCAAGAGAGAGACCAATCCGCGGGTTCATCAACAAAGGATTCCT 4080  
| | | | |  
QY 4082 CGGTCTCACTGAGGAGGACAGAGAGATCTCTGCGTGGGAGGCCAGTCTTAAAGCGG 4141  
| | | | |  
Db 4081 CGGTCTCACTGAGGAGGACAGAGAGATCTCTGCGTGGGAGGCCAGTCTTAAAGCGG 4140  
| | | | |  
QY 4142 GAGGAGAGCCTCCGCGCCCAACCGGCTCAACGAGACCTGACGAGGCGCTTAAAGAGCGAG 4201  
| | | | |  
Db 4141 GAGGAGAGCCTCCGCGCCCAACCGGCTCAACGAGACCTGACGAGGCGCTTAAAGAGCGAG 4200  
| | | | |  
QY 4202 GCCCTGGGCCCCCTGAAGCTGAAGCCGAGCCATGAGGGCTGTGTGGCAAGGTGAAGAG 4261  
| | | | |

Db 4201 GCCCTGGGCCCCCTGAAGCTGAAGCCGAGCCCATGAGGGGCTGTGTGGCAAGGTGAAGAG 4260  
| | | | |  
QY 4262 GCGGGCCGCTCATCATGAGATCCCGGCGAGGAGCTGCGGACACACGCCGAGCTGCC 4321  
| | | | |  
Db 4261 GCGGGCCGCTCATCATGAGATCCCGGCGAGGAGCTGCGGACACACGCCGAGCTGCC 4320  
| | | | |  
QY 4322 CTGGCCCCGCGCGCTCAAGAGAGGCTTCATCAACGAGGAGACCCGCTCAAGTACAGAC 4381  
| | | | |  
Db 4321 CTGGCCCCGCGCGCTCAAGAGAGGCTTCATCAACGAGGAGACCCGCTCAAGTACAGAC 4380  
| | | | |  
QY 4382 ACCGCGGGTCCACCACTGGGCTCCAAAAAGACAGAGTACGTCCTCCCTCATGCGGAGCCC 4441  
| | | | |  
Db 4381 ACCGCGGGTCCACCACTGGGCTCCAAAAAGACAGAGTACGTCCTCCCTCATGCGGAGCCC 4440  
| | | | |  
QY 4442 GCGCGGAGCTTCCCAACCCGCTGACCCGCTGATGATGAGCGGACGCGCGGACCTGGA 4501  
| | | | |  
Db 4441 GCGCGGAGCTTCCCAACCCGCTGACCCGCTGATGATGAGCGGACGCGCGGACCTGGA 4500  
| | | | |  
QY 4502 CGTGCCTGCTACGAGAGAGCTGAAGAGCCGCGCAGGAGACCGGACGAGCTCGGGGGGC 4561  
| | | | |  
Db 4501 CGTGCCTGCTACGAGAGAGCTGAAGAGCCGCGCAGGAGACCGGACGAGCTCGGGGGGC 4560  
| | | | |  
QY 4562 TCCATTTGCGCGCGCGCGCGCTCATTTGTGCTGAGCTGGGTAAAGCCCGCGAGAGCCCC 4621  
| | | | |  
Db 4561 TCCATTTGCGCGCGCGCGCGCTCATTTGTGCTGAGCTGGGTAAAGCCCGCGAGAGCCCC 4620  
| | | | |  
QY 4622 CTGACCTATGAGAGACACAGGGGACCTTTTGGCGGACACTTCCACAGAGTTTCGCGCTG 4681  
| | | | |  
Db 4621 CTGACCTATGAGAGACACAGGGGACCTTTTGGCGGACACTTCCACAGAGTTTCGCGCTG 4680  
| | | | |  
QY 4682 ACCATGCGGAGGCCACGCGCGCTGAGAGAGGAGAGGACCTTTTCGTCAAGAGGATCC 4741  
| | | | |  
Db 4681 ACCATGCGGAGGCCACGCGCGCTGAGAGAGGAGAGGACCTTTTCGTCAAGAGGATCC 4740  
| | | | |  
QY 4742 CAGGACCGAAAGCTGACGTCCAGCGCTGTGAGATTCGCCAAGTCCCGGACAGGACCGTG 4801  
| | | | |  
Db 4741 CAGGACCGAAAGCTGACGTCCAGCGCTGTGAGATTCGCCAAGTCCCGGACAGGACCGTG 4800  
| | | | |  
QY 4802 CCGGAGCACCAACCCCACTCGGCTGATGAGACCTTGCTTGGGGCGTGAAGTGGC 4861  
| | | | |  
Db 4801 CCGGAGCACCAACCCCACTCGGCTGATGAGACCTTGCTTGGGGCGTGAAGTGGC 4860  
| | | | |  
QY 4862 GTGACCTGTATTCGAGGACACATCCGCTGAGCTTGCACCCCACTTCATACCCGCGGC 4921  
| | | | |  
Db 4861 GTGACCTGTATTCGAGGACACATCCGCTGAGCTTGCACCCCACTTCATACCCGCGGC 4920  
| | | | |  
QY 4922 ATCCCTGTGAGAGCAGCGGCTGCTACTTACCTTGCCCGACACTTGGCCCCCAACCCAC 4981  
| | | | |  
Db 4921 ATCCCTGTGAGAGCAGCGGCTGCTACTTACCTTGCCCGACACTTGGCCCCCAACCCAC 4980  
| | | | |  
QY 4982 TACCGGCACTGTACCAACCCCTACTCATTCGCGGCTTACCCGACACAGGCGGCGCTGAG 5041  
| | | | |  
Db 4981 TACCGGCACTGTACCAACCCCTACTCATTCGCGGCTTACCCGACACAGGCGGCGCTGAG 5040  
| | | | |  
QY 5042 AACCGGACAGCATCATGATGATGATCACTTGCAGAGATGACACACAACGCGCC 5101  
| | | | |  
Db 5041 AACCGGACAGCATCATGATGATGATCACTTGCAGAGATGACACACAACGCGCC 5100  
| | | | |  
QY 5102 ACCGCGAGGCGCCAGCGAGCTGATGTGCTGAGGGGCTCTTGCCCGGAGTCTTCGCTG 5161  
| | | | |  
Db 5101 ACCGCGAGGCGCCAGCGAGCTGATGTGCTGAGGGGCTCTTGCCCGGAGTCTTCGCTG 5160  
| | | | |  
QY 5162 GCACTCAACTAGCGCTGCGGGTCCCGAGGACATCACTGACCTGTCCAGATGCAACACTG 5221  
| | | | |  
Db 5161 GCACTCAACTAGCGCTGCGGGTCCCGAGGACATCACTGACCTGTCCAGATGCAACACTG 5220  
| | | | |  
QY 5222 CCTGTGCTGTGTCGCCCGGACACAGGACCCGAGCCACCGGCAATGAGACCGCTTGTCTAC 5281  
| | | | |  
Db 5221 CCTGTGCTGTGTCGCCCGGACACAGGACCCGAGCCACCGGCAATGAGACCGCTTGTCTAC 5280  
| | | | |  
QY 5282 CTCCCAACGCGGCCCAAGCCTTTCAGAGCGGCGCACAGAGTCTCCCACTTCTCCAGGA 5341  
| | | | |



Db 5281 CTCGCCCGCGCCCGACGCCCTTCAGACAGCCGCGCAGACAGCTCCCACTCTCCCGAGGA 5340  
Qy 5342 GGTCAACAACACTTGAACAAACCAACCAACCAAGTCTCGTCCGAGCGGAGCGAGACCG 5401  
Db 5341 GGTCAACAACACTTGAACAAACCAACCAACCAAGTCTCGTCCGAGCGGAGCGAGACCG 5400  
Qy 5402 GATCGAGACCGGAGCCGGGATCGGAGCCGGAAAGTCCATCTCTCAAGTCCACACGACG 5461  
Db 5401 GATCGAGACCGGAGCCGGGATCGGAGCCGGAAAGTCCATCTCTCAAGTCCACACGACG 5460  
Qy 5462 GTGAGACACGACCCATCTGAGAACCTGTGTACAGACAGACAGCCGACAGCCGACG 5521  
Db 5461 GTGAGACACGACCCATCTGAGAACCTGTGTACAGACAGACAGCCGACAGCCGACG 5520  
Qy 5522 AGCGGCGGGGGGTGGGGGAGACAGACAGCCGCGCCCTCCCACTCCGATGCGCACAGAC 5581  
Db 5521 AGCGGCGGGGGGTGGGGGAGACAGACAGCCGCGCCCTCCCACTCCGATGCGCACAGAC 5580  
Qy 5582 TCGCCCATCTCCCTCGGAGCCAGATGCGCTTCAGACAGACCCAGTGTCTTCAACA 5641  
Db 5581 TCGCCCATCTCCCTCGGAGCCAGATGCGCTTCAGACAGACCCAGTGTGTCTTCAACA 5640  
Qy 5642 ACAAGCATGAAGGTATCATCACCGCTGTGAGCCAGACAGCCAGCTCTGAGGTCC 5701  
Db 5641 ACAAGCATGAAGGTATCATCACCGCTGTGAGCCAGACAGCCAGCTCTGAGGTCC 5700  
Qy 5702 ACCTCACCTCTCAACCGGTGGCCAGGTCGACATTCGCCATCTGCGCACCCATGCGCA 5761  
Db 5701 ACCTCACCTCTCAACCGGTGGCCAGGTCGACATTCGCCATCTGCGCACCCATGCGCA 5760  
Qy 5762 CTGGGCGGACCCCTCGATGGGGTCTACCCCTACCTCATGAGAGCCGCTTGTGCTGCCAAG 5821  
Db 5761 CTGGGCGGACCCCTCGATGGGGTCTACCCCTACCTCATGAGAGCCGCTTGTGCTGCCAAG 5820  
Qy 5822 GAGGCCCCCGGGGTGCGCCGCGCAGAGCGGCCCGCGACAGACACCGGCCATGCTTCTTC 5881  
Db 5821 GAGGCCCCCGGGGTGCGCCGCGCAGAGCGGCCCGCGACAGACACCGGCCATGCTTCTTC 5880  
Qy 5882 GCGAAGCCCCGCGGCTCCGGGCTGAGAGCCGCGCTCCCTCCCGCAGAAAGGCTCGGAG 5941  
Db 5881 GCGAAGCCCCGCGGCTCCGGGCTGAGAGCCGCGCTCCCTCCCGCAGAAAGGCTCGGAG 5940  
Qy 5942 CCCCGGCCCTAGTGTCTCTGTCTGTGACAGCCACCATGCGCCGACCCCTGCGAAG 6001  
Db 5941 CCCCGGCCCTAGTGTCTCTGTCTGTGACAGCCACCATGCGCCGACCCCTGCGAAG 6000  
Qy 6002 AACCTGCACTTCAACAGCGCCAGCCCGGACCCGCGCGCCCACTGCTCGGCTCGAGC 6061  
Db 6001 AACCTGCACTTCAACAGCGCCAGCCCGGACCCGCGCGCCCACTGCTCGGCTCGAGC 6060  
Qy 6062 CGGACCCGCGGAAAGACTMAAGTTAAACCTTTTCAATCAAGAACTGGAATCTCCGTTCT 6121  
Db 6061 CGGACCCGCGGAAAGACTMAAGTTAAACCTTTTCAATCAAGAACTGGAATCTCCGTTCT 6120  
Qy 6122 CTGGGTTACACAGCGCAGACTTACAGCCCGGAAAGGGGTGAGAGCCGCTGAGCCCTGTGAGC 6181  
Db 6121 CTGGGTTACACAGCGCAGACTTACAGCCCGGAAAGGGGTGAGAGCCGCTGAGCCCTGTGAGC 6180  
Qy 6182 TCACCCAGTCTGAACCAACGACAAAGGGGTCTCCCAAGCACCTGGAAGACTTGAACAGAC 6241  
Db 6181 TCACCCAGTCTGAACCAACGACAAAGGGGTCTCCCAAGCACCTGGAAGACTTGAACAGAC 6240  
Qy 6242 CACCTGAGAGGGGAGAGTGTGGGCCCAAGAGCAGAGCCCGGTGAAGTTTGGCGGAGAGGC 6301  
Db 6241 CACCTGAGAGGGGAGAGTGTGGGCCCAAGAGCAGAGCCCGGTGAAGTTTGGCGGAGAGGC 6300  
Qy 6302 GCGCACTTCCCAACACTGCGCGCTGTGAGAGCCAGCCCTGTCAAGCCGCTGTCTC 6361  
Db 6301 GCGCACTTCCCAACACTGCGCGCTGTGAGAGCCAGCCCTGTGTCAAGCCGCTGTCTC 6360  
Qy 6362 CAGACCGGCTCCAGGGGTCAAAGGTACACAGCGGGTGTCACTCTGCGCCAGCAATCACT 6421  
Db 6361 CAGACCGGCTCCAGGGGTCAAAGGTACACAGCGGGTGTCACTCTGCGCCAGCAATCACT 6420

Qy 6422 GAGGTATCAACAGGACTTACACCCGGGACCAACCCACAGAGCTCAGCGCACCCCTGCCC 6481  
Db 6421 GAGGTATCAACAGGACTTACACCCGGGACCAACCCACAGAGCTCAGCGCACCCCTGCCC 6480  
Qy 6482 GCGCCCTCTACTCTCTTCCCTGGGGGCAAGTGTCCGCTCTGGAACCTCGCGGCCACCC 6541  
Db 6481 GCGCCCTCTACTCTTCCCTGGGGGCAAGTGTCCGCTCTGGAACCTCGCGGCCACCC 6540  
Qy 6542 AGTGAACCTTACCTCCCGCCCGCGACATGAGTCCCGGCGCGGTATCCCGCCACAGC 6601  
Db 6541 AGTGAACCTTACCTCCCGCCCGCGACATGAGTCCCGGCGCGGTATCCCGCCACAGC 6600  
Qy 6602 GAAAGGGGCAAGAGTCTCAAGAGCCAAACAAGACTCGGTCTTGGAGTGTGAGAGC 6661  
Db 6601 GAAAGGGGCAAGAGTCTCAAGAGCCAAACAAGACTCGGTCTTGGAGTGTGAGAGC 6660  
Qy 6662 GGTATTGAACCTGTGTCCCAACCGAGGGCATGACGAGCCAGAGGACTTCCGAGTGCCT 6721  
Db 6661 GGTATTGAACCTGTGTCCCAACCGAGGGCATGACGAGCCAGAGGACTTCCGAGTGCCT 6720  
Qy 6722 GTGTACCCGCTGTGTACCGGGATGGGGAAACAAGAGCCCAAGACAGATGGGCTCAAG 6781  
Db 6721 GTGTACCCGCTGTGTACCGGGATGGGGAAACAAGAGCCCAAGACAGATGGGCTCAAG 6780  
Qy 6782 TCTCCAGGCAACACAGCCAGCGCCAGGCTTCTTCAAGCAAGTGAACGAGACACTCC 6841  
Db 6781 TCTCCAGGCAACACAGCCAGCGCCAGGCTTCTTCAAGCAAGTGAACGAGACACTCC 6840  
Qy 6842 GCCATGTCAAGTCCAAAGAGCAAGATCAACAAGACTGAACCCACACCGGAAT 6901  
Db 6841 GCCATGTCAAGTCCAAAGAGCAAGATCAACAAGACTGAACCCACACCGGAAT 6900  
Qy 6902 GAGCTGAATTAATATTCAGCCAGCTGGGAGCGGAATCTTCAATATGCCGCAATCAC 6961  
Db 6901 GAGCTGAATTAATATTCAGCCAGCTGGGAGCGGAATCTTCAATATGCCGCAATCAC 6960  
Qy 6962 GGAACAGGCTTATGACTATGAAGCCAGGCGGTGAGAAACATGSCAGACCAACATG 7021  
Db 6961 GGAACAGGCTTATGACTATGAAGCCAGGCGGTGAGAAACATGSCAGACCAACATG 7020  
Qy 7022 GGGCTGGAAGGCAATTAAGAAAGGCACTCATGGTTAAATATGACAGTGGAAAGTCC 7081  
Db 7021 GGGCTGGAAGGCAATTAAGAAAGGCACTCATGGTTAAATATGACAGTGGAAAGTCC 7080  
Qy 7082 CCGCGCTCAGGCGCAATCTTTTAAACCTTGAATGCAAGTGTCCGCTGCT 7141  
Db 7081 CCGCGCTCAGGCGCAATCTTTTAAACCTTGAATGCAAGTGTCCGCTGCT 7140  
Qy 7142 ATGCCATTAACGCTGTGAGAGGAGTGAACAACACTCAACCTCGCAGGGTGGCGG 7201  
Db 7141 ATGCCATTAACGCTGTGAGAGGAGTGAACAACACTCAACCTCGCAGGGTGGCGG 7200  
Qy 7202 GGGAAAGCCMAAGTCTTGTGAGAGCCAGCAGCCGAAAGCCAGTCCCGGCCCCGGGCG 7261  
Db 7201 GGGAAAGCCMAAGTCTTGTGAGAGCCAGCAGCCGAAAGCCAGTCCCGGCCCCGGGCG 7260  
Qy 7262 CTGGCATCTGGGGAACCGGCACCTCTGTCTCTCAATGTCATCTCGAGGGAAGTGTGAAC 7321  
Db 7261 CTGGCATCTGGGGAACCGGCACCTCTGTCTCTCAATGTCATCTCGAGGGAAGTGTGAAC 7320  
Qy 7322 GCGCGGAGCGGCTCAACCAACGCGTGTGGGAGAGCAAGGCGCTCGGCGCAGGTTCAAG 7381  
Db 7321 GCGCGGAGCGGCTCAACCAACGCGTGTGGGAGAGCAAGGCGCTCGGCGCAGGTTCAAG 7380  
Qy 7382 CCATTCCTCTACACCCCTGTATCATGCGGCTGACAGCGGGTGTATGAGCTTCCCAACC 7441  
Db 7381 CCATTCCTCTACACCCCTGTATCATGCGGCTGACAGCGGGTGTATGAGCTTCCCAACC 7440  
Qy 7442 CCAACGGGCTTCCCGCGGAGCGGGGCCCTGTGCTGGCCCCCAACAAGCTGTGGAGCAG 7501  
Db 7441 CCAACGGGCTTCCCGCGGAGCGGGGCCCTGTGCTGGCCCCCAACAAGCTGTGGAGCAG 7500

QY 7502 GAGCCCAAGCCACTGCTCTGCGAGTACGAGACACTCTCCGACGCGAGTGA 7555  
DB 7501 GAGCCCAAGCCACTGCTCTGCGAGTACGAGACACTCTCCGACGCGAGTGA 7554

RESULT 14

ID ADL13812 standard; DNA; 7524 BP.

ADL13812;

06-MAY-2004 (first entry)

Osteoarthritis-associated polymorphic nucleotide #344.

ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;  
joint space narrowing; osteophyte development; joint pain;  
osteoarthritis; SNP; single nucleotide polymorphism.

Homo sapiens.

MO200305416-A2.

03-JUL-2003.

19-DEC-2002; 2002MO-US041225.

20-DEC-2001; 2001US-0342603P.

(INCY-) INCYTE GENOMICS INC.

Jones KA, Schafer A;

WPI; 2003-559141/52.

Determining susceptibility of an individual to joint space narrowing,  
osteophyte development and/or joint pain comprises identifying whether  
the individual has at least one polymorphism in a polynucleotide encoding  
a protein.

Disclosure: SEQ ID NO 344; 297bp; English.

The invention relates to a method of determining susceptibility of an  
individual to joint space narrowing and/or osteophyte development and/or  
joint pain comprising identifying whether the individual has at least one  
polymorphism in a polynucleotide encoding at least one of the protein  
listed in the specification. The methods, composition and agent are  
useful for modulating the susceptibility of an individual to joint space  
narrowing and/or osteophyte development and/or joint pain that is  
associated with a disease, preferably osteoarthritis. The cell line and  
the non-human animal are useful for screening for an agent for diagnosing  
an individual having susceptibility to joint space narrowing and/or  
osteophyte development and/or joint pain. This sequence corresponds to  
the polynucleotide encoding a protein listed in the specification. (Note:  
The sequence data for this patent did not form part of the printed  
specification but was obtained in electronic format directly from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences).

Sequence 7524 BP; 1636 A; 2728 C; 2213 G; 947 T; 0 U; 0 Other;

Query Match 85.7%; Score 7337.6; DB 10; Length 7524;

Best Local Similarity 98.5%; Pred. No. 0;

Matches 7466; Conservative 0; Mismatches 34; Indels 78; Gaps 3;

2 ATGTGGGCTCCACAGCTTGTGGACAGATGAGAGGCGACGAGCCCGCTACCG 61

1 ATGTGGGCTCCACAGCTTGTGGACAGATGAGAGGCGACGAGCCCGCTACCG 60

62 CCCACAGACCTTTCTTACCCAGTGCAGATCGCCGGACGACACGAGCGTGGCTCTG 121

61 CCCACAGACCTTTCTTACCCAGTGCAGATCGCCGGACGACACGAGCGTGGCTCTG 120

122 GAGTACGACGACCACTCCCGCGACTATGCTCCACTGTGCGCGGGCTCATATCCAG 181

DB 121 GAGTACGACGACCACTCCCGCGACTATGCTCCACTGTGCGCGGGCTCATATCCAG 180

QY 182 CCCACAGACGAGGAGGCTCCCTGCTGTGTAGATTCCAGCCCGGAAATGAACGGTCCAG 241

DB 181 CCCACAGGCGGAGGAGGCTCCCTGCTGTGTAGATTCCAGCCCGGAAATGAACGGTCCAG 240

QY 242 GAGTCCACTGTCGGCGCAGAGTCCCATATCTGCTCCGAGCTGGGAGTCCAGATG 301

DB 241 GAGTCCACTGTCGGCGCAGAGTCCCATATCTGCTCCGAGCTGGGAGTCCAGATG 300

QY 302 GAGTTCAATTGAAAGAAAGACCGCTTCGCTAGAGCTGCTGCAACCCCTGCTGGAGACG 361

DB 301 GAGTTCAATTGAAAGAAAGACCGCTTCGCTAGAGCTGCTGCAACCCCTGCTGGAGACG 360

QY 362 TCACCCCTGCTGGCGCAGCGGCGAGCTCGGGATCTGAAGACCTCACAGGACCGTAGC 421

DB 361 TCACCCCTGCTGGCGCAGCGGCGAGCTCGGGATCTGAAGACCTCACAGGACCGTAGC 420

QY 422 CTGACGGGCAAGCTTGAAGACCGGTGTCTCCCGCAGCGCCCGCAGACCTGAGCTG 481

DB 421 CTGACGGGCAAGCTTGAAGACCGGTGTCTCCCGCAGCGCCCGCAGACCTGAGCTG 480

QY 482 GAGCTGTGTCGGCGCAGCGGCTGTCCAGAGGAGAGCTGATCCAGAACTAGACCGCTGAC 541

DB 481 GAGCTGTGTCGGCGCAGCGGCTGTCCAGAGGAGAGCTGATCCAGAACTAGACCGCTGAC 540

QY 542 CGAGAGATCACCATGTGTAGACAGACAGATCTCTAAGCTGAAGAAAGAACAGACAGCTG 601

DB 541 CGAGAGATCACCATGTGTAGACAGACAGATCTCTAAGCTGAAGAAAGAACAGACAGCTG 600

QY 602 GAGGAGAGGCTGTGCCAAGCGCGCCGAGGCTTGAGAAAGCCCTGTGTCAACGCGCCCATCAG 661

DB 601 GAGGAGAGGCTGTGCCAAGCGCGCCGAGGCTTGAGAAAGCCCTGTGTCAACGCGCCCATCAG 660

QY 662 TCGAAGCACCCGACGCTGTGTCAAGATCATCTACCAACGGAACCCGGAAGAAGCTGAAGCT 721

DB 661 TCGAAGCACCCGACGCTGTGTCAAGATCATCTACCAACGGAACCCGGAAGAAGCTGAAGCT 720

QY 722 GCACATCGGATTTCTGGAAGGCTGGGGGCCCGAGGTGAGTGCCTGTACACCAAGCCG 781

DB 721 GCACATCGGATTTCTGGAAGGCTGGGGGCCCGAGGTGAGTGCCTGTACACCAAGCCG 780

QY 782 TCCGACACCCGCGAGATCATGAGAACTCAAAATTAACAGCGCATCGGAGAGACTTA 841

DB 781 TCCGACACCCGCGAGATCATGAGAACTCAAAATTAACAGCGCATCGGAGAGACTTA 840

QY 842 ATCTTGTACTTCAAGAGAGGAAATCAAGCTGTGGAAACAAATGGAACAGAAATTTCCAG 901

DB 841 ATCTTGTACTTCAAGAGAGGAAATCAAGCTGTGGAAACAAATGGAACAGAAATTTCCAG 900

QY 902 CGCTATGACAGACTCATGAGAGGCTTGGAAAAAAGGTGAGCGCATGAAAAACAACCG 961

DB 901 CGCTATGACAGACTCATGAGAGGCTTGGAAAAAAGGTGAGCGCATGAAAAACAACCG 960

QY 962 CGCGGCGGGCGCAAGAGAGCAAGGTGCGAGTACTACGAAAAAGCAAGTCTCTGAGATC 1021

DB 961 CGCGGCGGGCGCAAGAGAGCAAGGTGCGAGTACTACGAAAAAGCAAGTCTCTGAGATC 1020

QY 1022 CGCAAGACGCGCGAGCTTCAAGAGGCGCATGCAAGACAGGTGGGCGACGGGCGAGTGG 1081

DB 1021 CGCAAGACGCGCGAGCTTCAAGAGGCGCATGCAAGACAGGTGGGCGACGGGCGAGTGG 1077

QY 1082 CTGTTCATGTGGCGCGCGCGCGAGGAGCAGAGGTGTGAGAGATCATGATGAGCTCTCA 1141

DB 1078 CTGTTCATGTGGCGCGCGCGCGAGGAGCAGAGGTGTGAGAGATCATGATGAGCTCTCA 1137

QY 1142 GAGCAGAGAACTTGAAGACAGATGCGCAAGTGTGCGCTGATCCGCGCATGTGTAC 1201

DB 1138 GAGCAGAGAACTTGAAGACAGATGCGCAAGTGTGCGCTGATCCGCGCATGTGTAC 1197

QY 1202 GAGCTGACGACGAGCGATCAAGTTATCAACATGAACGGGCTTATGGCCGACCCCATG 1261

Db 1198 GACGCTGACGACGACGACATCAATTCAATCAATGACGGGCTTATGCGGACCCCATG 1257  
Qy 1262 AAGGTGTACAAAGACCCGACGATCATGAACATGTGAGTGAAGAGAGAACCTTC 1321  
Db 1258 AAGGTGTACAAAGACCCGACGATCATGAACATGTGAGTGAAGAGAGAACCTTC 1317  
Qy 1322 CGGAGAAATTCAATGACGATCCCAAGAACTTGGCTGATGCGATCATTTCTGAGAG 1381  
Db 1318 CGGAGAAATTCAATGACGATCCCAAGAACTTGGCTGATGCGATCATTTCTGAGAG 1377  
Qy 1382 AAGACATGCTGATGATGCTGCTCTATTTACTTACTTACCTTACAGAAATGAAGACTTAAG 1441  
Db 1378 AAGACATGCTGATGATGCTGCTCTATTTACTTACTTACCTTACAGAAATGAAGACTTAAG 1437  
Qy 1442 AGCTGTGTAGA CCGAGCTATGCGGCGCCGCGAAGAGCAGAGAGCAACAGCAGCAG 1501  
Db 1438 AGCTGTGTAGA CCGAGCTATGCGGCGCCGCGAAGAGCAGAGAGCAACAGCAGCAG 1497  
Qy 1502 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1561  
Db 1498 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1557  
Qy 1562 GATGAGAAAG 1621  
Db 1558 GATGAGAAAG 1617  
Qy 1622 GACAGGAAGACCTCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681  
Db 1618 GACAGGAAGACCTCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1677  
Qy 1682 AAGAGAGCTGTGGCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1741  
Db 1678 AAGAGAGCTGTGGCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1737  
Qy 1742 GCGATCAACCCGCTCAATGAGTGAAGGCAACAGGAGAGAGAGAGAGAGAGAGAGAG 1801  
Db 1738 GCGATCAACCCGCTCAATGAGTGAAGGCAACAGGAGAGAGAGAGAGAGAGAGAGAG 1797  
Qy 1802 AGCGCGAGCTGGCTCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1861  
Db 1798 AGCGCGAGCTGGCTCCCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1857  
Qy 1862 GAAAGAGCAGAAAGAGCTCTCTGAGAACAGGCGGCACTGGTCGCGCATTCGCGCGATG 1921  
Db 1858 GAAAGAGCAGAAAGAGCTCTCTGAGAACAGGCGGCACTGGTCGCGCATTCGCGCGATG 1917  
Qy 1922 GTGGGCTCCAAAGCTGTGTGCGAGTGTAAAGAACTTCTACTTCAACTTCAAGAAAGAGCAG 1981  
Db 1918 GTGGGCTCCAAAGCTGTGTGCGAGTGTAAAGAACTTCTACTTCAACTTCAAGAAAGAGCAG 1977  
Qy 1982 AACCTTGATGAGATCTTTCAGACGACAGCTGAGATGAGAGAGAGAGAGAGAGAGAGAG 2041  
Db 1978 AACCTTGATGAGATCTTTCAGACGACAGCTGAGATGAGAGAGAGAGAGAGAGAGAGAG 2037  
Qy 2042 AGGAGAAAGAGAAAGAGCGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2101  
Db 2038 AGGAGAAAGAGAAAGAGCGCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2097  
Qy 2102 GAGGATGAGAGATGAGAGAGCGTTCGAGCGTGAAGCGAGAAATGAGAGAGAGATGAGAGAGAG 2161  
Db 2098 GAGGATGAGAGATGAGAGAGCGTTCGAGCGTGAAGCGAGAAATGAGAGAGAGATGAGAGAGAG 2157  
Qy 2162 GGTGAAGCTTTCATAGCTCTCTGAGAAATGAGTGTGCCAGAGGGGAAATGCAAGTGGCCCAAGCC 2221  
Db 2158 GGTGA-----AGCC 2166  
Qy 2222 ACTGTCAACAAAGCTCAGACACCGAGAGCATCCCTCTCTCACTGAGCGCGCAAG 2281  
Db 2167 ACTGTCAACAAAGCTCAGACACCGAGAGCATCCCTCTCTCACTGAGCGCGCAAG 2226  
Qy 2282 GACACAGGCGAGAAATGGGCTCAAGGCCCGCAGCCACCTTGAGCGCGCAGCGGCGCACCCCA 2341  
Db 2227 GACACAGGCGAGAAATGGGCTCAAGGCCCGCAGCCACCTTGAGCGCGCAGCGGCGCACCCCA 2286

Qy 2342 GGCCCAACCCCAACCCAGAGAGACATCCGGGCGCCCATTTGAGCCACCCCGGCTCT 2401  
Db 2287 GGCCCAACCCCAACCCAGAGAGACATCCGGGCGCCCATTTGAGCTCAACCCCGGCTCT 2346  
Qy 2402 GAAGCCACCGAGAGCCCTTACCGCTCCCAAGAGACCCCACTGCGCTCTGACCTCTCTCT 2461  
Db 2347 GAAGCCACCTTACCGCTTACCGCTCCCAAGAGACCCCACTTCTCTTCACTCTCTCTCT 2406  
Qy 2462 GTGTCTCCCAAG 2521  
Db 2407 GTGTCTCCCAAG 2466  
Qy 2522 GAGAGACAG 2581  
Db 2467 GAGAGACAG 2526  
Qy 2582 CCGCTCAAG 2641  
Db 2527 CCGCTCAAG 2586  
Qy 2642 GAGGCGCTGAG 2701  
Db 2587 GAGGCGCTGAG 2646  
Qy 2702 GGCAGGGCCACACCTGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2761  
Db 2647 GGCAGGGCCACACCTGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2706  
Qy 2762 TGCAGTGCAGAGAGAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2821  
Db 2707 TGCAGTGCAGAGAGTGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2766  
Qy 2822 AAGCCAGCTCTCTCACTCCGACCTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2881  
Db 2767 AAGCCAGCTCTCTCACTCCGACCTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2826  
Qy 2882 CTGGAAGTGAAG 2941  
Db 2827 CTGGAAGTGAAG 2886  
Qy 2942 GTTCATGAG 3001  
Db 2887 GTTCATGAG 2946  
Qy 3002 CCAAGGCAAAACCTGAG 3061  
Db 2947 CCAAGGCAAAACCTGAG 3006  
Qy 3062 GGCAG 3100  
Db 3007 GGCAG 3066  
Qy 3101 ---GCCCTTCAG 3157  
Db 3067 CCAAGCTTCGAG 3126  
Qy 3158 CTGCGCTTCCTGAG 3217  
Db 3127 CTGCGCTTCCTGAG 3186  
Qy 3218 TCAGGCTTCCTGAG 3277  
Db 3187 TCAGGCTTCCTGAG 3246  
Qy 3278 CGGCGGCTCTGAG 3337  
Db 3247 CGGCGGCTCTGAG 3306  
Qy 3338 AAGCAACCCAG 3397  
Db 3307 AAGCAACCCAG 3366

QY	3398	TTCCACGCTCCGCTACTCAAGACATGACCAAGGCCCCCGGTGGGCGCTGTCAACATGGGGCTG	3457
Dp	3367	CTCAACGCTCCGTAAGTCAAGACATGACCAAGGCCCCCGGTGGGCGCTGTCAACATGGGGCTG	3426
QY	3458	CCCTGTCCATGAGACCCCAAAAAGCTGGCACTCTTCAGCCGAGTGAAGCAGAGACAGCTG	3517
Dp	3427	CCCTGTCCATGAGACCCCAAAAAGCTGGCACTCTTCAGCCGAGTGAAGCAGAGACAGCTG	3486
QY	3518	TTCCCAACGGGGGCCAAGCTGTGGGCCACCCGAGACCTTGGGGGTGCCCAAGCCCAAGAGCGC	3577
Dp	3487	TTCCCAACGGGGGCCAAGCTGTGGGCCACCCGAGAGCTTGGGGGTGCCCAAGCCCAAGAGCGC	3546
QY	3578	TTCCGTGCTGAAGAGGAGACAGCTGTGGGGCTCAATTCGCGGGCCGAGATCAACAAAGGGCAT	3637
Dp	3547	TTCCGTGCTGAAGAGGAGACAGCTGTGGGGCTCAATTCGCGGGCCGAGATCAACAAAGGGCAT	3606
QY	3638	CCCAAGCACACGGGTGGCCCTCCGAGACAGCGCCATCAATACCCGGCTCCATCAACCACGAGC	3697
Dp	3607	CCCAAGCACACGGGTGGCCCTCCGAGACAGCGCCATCAATACCCGGCTCCATCAACCACGAGC	3666
QY	3698	ACGCCAGCTGACGCTCTGTACAGAGGACCAATCAACAGAGATCATGTGGCAGAGACAGCCCG	3757
Dp	3667	ACGCCAGCTGACGCTCTGTACAGAGGACCAATCAACAGAGATCATGTGGCAGAGACAGCCCG	3726
QY	3758	AGTGGCTTTGACCCCGCGCCGGAGAGACAGCCCTGCGCAAGGGGCCAGTCAATTAAGAAAGGC	3817
Dp	3727	AGTGGCTTTGACCCCGCGCCGGAGAGACAGCCCTGCGCAAGGGGCCAGTCAATTAAGAAAGGC	3786
QY	3818	AAGAAGGGCCACAGCTTGTCTCTATGAGAGGTGAGCATGTCTGTGACCCAGTGTCTCAAGAG	3877
Dp	3787	AAGAAGGGCCACAGCTTGTCTCTATGAGAGGTGAGCATGTCTGTGACCCAGTGTCTCAAGAG	3846
QY	3878	GACGCGCAGAGACAGCTCAGAACCCCTCCCATGAGACGCGCCGCCCAAGCGCACCTATGAC	3937
Dp	3847	GACGCGCAGAGACAGCTCAGAACCCCTCCCATGAGACGCGCCGCCCAAGCGCACCTATGAC	3906
QY	3938	ATTGATGAGAGGGCCCGCTGGGGGAGAGGCCATCTCTCAGACACATCGAAGGTCTATGGGC	3997
Dp	3907	ATTGATGAGAGGGCCCGCTGGGGGAGAGGCCATCTCTCAGACACATCGAAGGTCTATGGGC	3966
QY	3998	CGTCCCATCTCCCGCCGAGCGAGCAGACAGCCGCCACCACTCAAGAGAGCAGCACATCTCCG	4057
Dp	3967	CGTCCCATCTCCCGCCGAGCGAGCAGACAGCCGCCACCACTCAAGAGAGCAGCACATCTCCG	4026
QY	4058	GGGTTCATCACAACAAGGATCTCTGGTCTTACGTGAGAGGACAGAGAGACTTACTGTGCT	4117
Dp	4027	GGGTTCATCACAACAAGGATCTCTGGTCTTACGTGAGAGGACAGAGAGACTTACTGTGCT	4086
QY	4118	CGGAGAGCGAGACTCTCTTAAACCGGGAGGGGACAGCTCCGCCCCCAACCGGCTCTACGGGAC	4177
Dp	4087	CGGAGAGCGAGACTCTCTTAAACCGGGAGGGGACAGCTCCGCCCCCAACCGGCTCTACGGGAC	4146
QY	4178	CTGACCCGAGGCGCTTCAAGACGCAAGGCGCTTGGAGCCCTCTGAAGCTGAAGCCGAGCCATGAG	4237
Dp	4147	CTGACCCGAGGCGCTTCAAGACGCAAGGCGCTTGGAGCCCTCTGAAGCTGAAGCCGAGCCATGAG	4206
QY	4238	GGCTGTGTGGCCACGCTGAAGAGAGGCGGGCCGCTTCATTCATGAGATCCGCGGAGAGAG	4297
Dp	4207	GGCTGTGTGGCCACGCTGAAGAGAGGCGGGCCGCTTCATTCATGAGATCCGCGGAGAGAG	4266
QY	4298	CTGGGGCACAACGCCCGAGCTGCCCCCTGGCCCCCGGAGCGCTCAACATGAGATCCCGCGAGAGAG	4357
Dp	4267	CTGGGGCACAACGCCCGAGCTGCCCCCTGGCCCCCGGAGCGCTCAACATGAGATCCCGCGAGAGAG	4326
QY	4358	CAGGGCACCCCGGCTCAAGTACAGACACCGGGCGGTTCACATGGCTCCAAAACACAGAC	4417
Dp	4327	CAGGGCACCCCGGCTCAAGTACAGACACCGGGCGGTTCACATGGCTCCAAAACACAGAC	4386
QY	4418	GTAAGCTCTCTCATTCGAGAGCCCGGACGGAGAGTTCACACCCGCTGACCCCGCTGATGTG	4477
Dp	4387	GTAAGCTCTCTCATTCGAGAGCCCGGACGGAGAGTTCACACCCGCTGACCCCGCTGATGTG	4446
QY	4478	ATGGCCGACGCGCCGGGACCTGGAACCTGCTACGAGGAGAGCGCTTGAAAGCCCGGAC	4537

Db	4447	ATGGCCCAACGCCCCGGGACCTGGAACTGTGCTCTGCTACAGAGAGAGCTGGAAGGCCGGCCCA	4506
Oy	4538	GGAGCCGCGACAGACTCGG9999GCTCCATTGGCGCGCGCGCCCCCGGTCATTGTGCTTGAG	4597
Db	4507	GGGACCCCGACAGACTCGG9999GCTCCATTGGCGCGCGCGCCCCCGGTCATTGTGCTTGAG	4566
Oy	4598	CTGGGTAAAGCCCGGGCAGAGCCCCCTGTACCTATTAGGAGACACCGGGGGACACCTTTGGCGGGC	4657
Db	4567	CTGGGTAAAGCCCGGGCAGAGCCCCCTGTACCTATTAGGAGACACCGGGGGACACCTTTGGCGGGC	4626
Oy	4658	CACCTCCCAAGAGGTTGGCCCGGTGACCATGCGGGAGGCCACCGCGCGCCCTTGACAGAGAGGC	4717
Db	4627	CACCTCCCAAGAGGTTGGCCCGGTGACCATGCGGGAGGCCACCGCGCGCCCTTGACAGAGAGGC	4686
Oy	4718	AGCCTTTGCTTCAGCAAGGCATCCAGAGACCGAAAGCTGACGTGACGCTTCGTGAGATC	4777
Db	4687	AGCCTTTGCTTCAGCAAGGCATCCAGAGACCGAAAGCTGACGTGACGCTTCGTGAGATC	4746
Oy	4778	GCCAAATCCCCCGCACAGACACGTGCCCCGAGACACACCCACACCCCATTTGGCCCTTAGAG	4837
Db	4747	GCCAAATCCCCCGCACAGACACGTGCCCCGAGACACACCCACACCCCATTTGGCCCTTAGAG	4806
Oy	4838	CACCTGCTTCGGGGCGTGAATGGGTGAGACTGTGATGTGACGACACATCCCTTGACCTTC	4897
Db	4807	CACCTGCTTCGGGGCGTGAATGGGTGAGACTGTGATGTGACGACACATCCCTTGACCTTC	4866
Oy	4898	GACCCCACTTCATATCCCGGGGATCCCTGTGAGCGAGACCGGTGCTTACTACTTGCCCC	4957
Db	4867	GACCCCACTTCATATCCCGGGGATCCCTGTGAGCGAGACCGGTGCTTACTACTTGCCCC	4926
Oy	4958	CGAACACTGGCCCCCAACCCCACTTACCCGACACTTGATACCCACCTTACCTATCCCGGGC	5017
Db	4927	CGAACACTGGCCCCCAACCCCACTTACCCGACACTTGATACCCACCTTACCTATCCCGGGC	4986
Oy	5018	TACCCCGACACGGGCGGCGCTGGAGAACCGGAGACCATCATCTCAATGACTCATCACTCG	5077
Db	4987	TACCCCGACACGGGCGGCGCTGGAGAACCGGAGACCATCATCTCAATGACTCATCACTCG	5046
Oy	5078	CAGCAGATGACACCAACACGGGCCACCGGCATGAGCCGAGAGCTGATATGTTAGGG9GC	5137
Db	5047	CAGCAGATGACACCAACACGGGCCACCGGCATGAGCCGAGAGCTGATATGTTAGGG9GC	5106
Oy	5138	CTCTGCGCCCCGCGAGTCTTGCTGTGGACTCATCTACGCTGTGGGTCCCCCGAGGACATCATC	5197
Db	5107	CTCTGCGCCCCGCGAGTCTTGCTGTGGACTCATCTACGCTGTGGGTCCCCCGAGGACATCATC	5166
Oy	5198	GACCTGTCCCAAGTGGACACACTGCGCTGTGCTGTCGCCCGACACACCGAGCCACCCGACGC	5257
Db	5167	GACCTGTCCCAAGTGGACACACTGCGCTGTGCTGTCGCCCGACACACCGAGCCACCCGACGC	5226
Oy	5258	ACCGCATATGACCGGCTTGGCTTACTTCCCAACCGGCGCCGAGCCCTTACAGAGCCGCGCAC	5317
Db	5227	ACCGCATATGACCGGCTTGGCTTACTTCCCAACCGGCGCCGAGCCCTTACAGAGCCGCGCAC	5286
Oy	5318	AGCAGCTTCCCACTCTTCCCAAGAGTTCACACACTTGTACAAACCAACCAACCAACGATCC	5377
Db	5287	AGCAGCTTCCCACTCTTCCCAAGAGTTCACACACTTGTACAAACCAACCAACCAACGATCC	5346
Oy	5378	TCTGTCCAGAGGGGAGCGAGACCGGGATCGAGAGCGGGATCGGGAGCGGGAAAG	5437
Db	5347	TCTGTCCAGAGGGGAGCGAGACCGGGATCGAGAGCGGGATCGGGAGCGGGAAAG	5406
Oy	5438	TCCATCTCTCAAGTTCACCAAGACGGTGGACACGCACTTGTGAGACCTGTGTACAGAG	5497
Db	5407	TCCATCTCTCAAGTTCACCAAGACGGTGGACACGCACTTGTGAGACCTGTGTACAGAG	5466
Oy	5498	CAGAGCAGCGGACAGCGGACGAGCGCGCGGGGTGG9999CAGACAGACGCGCCCGCC	5557
Db	5467	CAGAGCAGCGGACAGCGGACGAGCGCGCGGGGTGG9999CAGACAGCGCGCCCGCC	5526
Oy	5558	TCCCATCTCCCATGCTCCACAGCACTGGCCCATCTTCCCTTGGAACCGAGATGCCCTTCAG	5617



KM tissue typing; cancer.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FT CDS 1..7524  
FT /tag= a  
FT /partial  
FT /product= "SMRTE"  
FT /note= "No stop codon given"  
XX US2003027137-A1.  
XX 06-FEB-2003.  
XX 27-MAR-2001; 2001US-00819104.  
XX 29-MAR-2000; 2000US-0193138P.  
XX (CHEN/) CHEN J D.  
XX Chen JD;  
XX WPI: 2003-466139/44.  
XX P-PSDB; ABU61812.  
XX New SMRTE proteins and nucleic acids, useful in gene therapy, predictive  
PT medicine, therapeutic or prophylactic treatment, chromosome mapping,  
PT tissue typing and in forensic biology.  
XX Claim 2; Page 48-56; 90pp; English.  
XX The invention relates to an isolated SMRTE nucleic acid molecule. The  
CC nucleic acids are useful in gene therapy, as hybridisation probes for  
CC identifying SMRTE-encoding nucleic acid molecules and as primers for  
CC amplifying of SMRTE nucleic acid molecules. The polypeptides are useful  
CC as immunogens to raise anti-SMRTE antibodies. The SMRTE molecules are  
CC useful as targets for discovering and developing modulating agents to  
CC regulate a variety of cellular processes, in screening assays, in  
CC predictive medicine, in therapeutic or prophylactic treatment, in  
CC chromosome mapping, tissue typing and in forensic identification of a  
CC biological sample. Modulators of SMRTE are useful for treating or  
CC preventing a condition associated with aberrant SMRTE protein or nucleic  
CC acid expression or activity, such as cancer. The present sequence  
CC represents the human nuclear receptor corepressor SMRTE coding region  
CC cDNA  
XX  
SQ Sequence 7521 BP, 1635 A; 2728 C; 2212 G; 946 T; 0 U; 0 Other;  
Query Match 85.6%; Score 7334.6; DB 8; Length 7521;  
Best Local Similarity 98.5%; Pred. No. 0;  
Matches 7463; Conservative 0; Mismatches 34; Indels 78; Gaps 3;  
QY 2 ATGTGGGCTCCACACAGCTTGTGGCAGACGTGAGGGCCACTGAGCCCGTACCCG 61  
DB 1 ATGTGGGCTCCACACAGCTTGTGGCAGACGTGAGGGCCACTGAGCCCGTACCCG 60  
QY 62 CCCACAGCCTTCTTCCACAGTGAATCGCCGGAGCCACAGGAGTGGGCTCCCG 121  
DB 61 CCCACAGCCTTCTTCCACAGTGAATCGCCGGAGCCACAGGAGTGGGCTCCCG 120  
QY 122 GAGTACGACACCACTCCCGGACATATGCTCCACCTGTCCGCGGGCTCATCATCCAG 181  
DB 121 GAGTACGACACCACTCCCGGACATATGCTCCACCTGTCCGCGGGCTCATCATCCAG 180  
QY 182 CCCACGGCGGAGGCTCTCTGTCTGTGAGTTCCAGCCCGGAGATGAACGGTCCAG 241  
DB 181 CCCACGGCGGAGGCTCTCTGTCTGTGAGTTCCAGCCCGGAGATGAACGGTCCAG 240  
QY 242 GAGCTCACTGGCGGACAGATCCACTATACCTGCGCGAGCTGGGAGGACAGAGATG 301  
DB 241 GAGCTCACTGGCGGACAGATCCACTATACCTGCGCGAGCTGGGAGGACAGAGATG 300

QY 302 GAGTTATTGAAGCAAGCGCTTCGAGTAGAGCTGTGCTTGAACCCCTGTGCGACCG 361  
DB 301 GAGTTATTGAAGCAAGCGCTTCGAGTAGAGCTGTGCTTGAACCCCTGTGCGACCG 360  
QY 362 TCACCCCTGTGGCCACAGGCGCAGCTGCGGGATCTGAAAGCTTCACGAAGACCTGAGC 421  
DB 361 TCACCCCTGTGGCCACAGGCGCAGCTGCGGGATCTGAAAGCTTCACGAAGACCTGAGC 420  
QY 422 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCAGCCCCCGGACACATGACCTTGAGCTG 481  
DB 421 CTGACGGGCAAGCTGGAACCGGTGTCTCCCCAGCCCCCGGACACATGACCTTGAGCTG 480  
QY 482 GAGCTGTGCGCCACAGGCTGTCCAAAGGAGAGCTGATCCAGAACATGACCGCTGAC 541  
DB 481 GAGCTGTGCGCCACAGGCTGTCCAAAGGAGAGCTGATCCAGAACATGACCGCTGAC 540  
QY 542 CGAGAGATCACCATGTGTGAGCAGAGATCTCTTAAGCTGAAGAAAGAACAGCAACGCTG 601  
DB 541 CGAGAGATCACCATGTGTGAGCAGAGATCTCTTAAGCTGAAGAAAGAACAGCAACGCTG 600  
QY 602 GAGAGAGAGCTGCGCAAGCGCTGCGAGCGCTGAGAAAGCCGCTGTCAACCGCCCATCGAG 661  
DB 601 GAGAGAGAGCTGCGCAAGCGCTGCGAGCGCTGAGAAAGCCGCTGTCAACCGCCCATCGAG 660  
QY 662 TCGAAGCACCGCAGCTGTGCGAATCATCTACGACGAAACCGGAGAGAGCTGAGCT 721  
DB 661 TCGAAGCACCGCAGCTGTGCGAATCATCTACGACGAAACCGGAGAGAGCTGAGCT 720  
QY 722 GCACATCGGATTCGGAAGGCTGGGGCCCGAGGTGAGGCTGCGCTGTACACCAAGGCC 781  
DB 721 GCACATCGGATTCGGAAGGCTGGGGCCCGAGGTGAGGCTGCGCTGTACACCAAGGCC 780  
QY 782 TCCGACACCCGCGAGTATCATGAGAACATCAAAATTAACAGCGCATCGGAGAGAGCTA 841  
DB 781 TCCGACACCCGCGAGTATCATGAGAACATCAAAATTAACAGCGCATCGGAGAGAGCTA 840  
QY 842 ATCTTTGATCTTAAGAGAGAGAAATCAAGCTCGGAAACATGGAAGCAGAGTTCTCGAG 901  
DB 841 ATCTTTGATCTTAAGAGAGAGAAATCAAGCTCGGAAACATGGAAGCAGAGTTCTCGAG 900  
QY 902 CGCATATGACGATCATGAGAGGCTTGGAAAAAAGGTGAGCGCATGGAAGAACACCG 961  
DB 901 CGCATATGACGATCATGAGAGGCTTGGAAAAAAGGTGAGCGCATGGAAGAACACCG 960  
QY 962 CGCGGCGGGCCAAAGAGAGCAAGGTGCGGAGTACTGAAAAAGCAATTCCTGAGATC 1021  
DB 961 CGCGGCGGGCCAAAGAGAGCAAGGTGCGGAGTACTGAAAAAGCAATTCCTGAGATC 1020  
QY 1022 CGAAAGCAGCGGAGCTGCAAGCGCATGCAAGAGCAGGCTGAGCGGCGCATGCGG 1081  
DB 1021 CGAAAGCAGCGGAGCTGCAAGCGCATGCAAGAGCAGGCTGAGCGGCGCATGCGG 1077  
QY 1082 CTGTTCATGTGGGCGCCCGGAGGAGCAAGAGTGTGAGAGATCATGATGAGGCTCTCA 1141  
DB 1078 CTGTTCATGTGGGCGCCCGGAGGAGCAAGAGTGTGAGAGATCATGATGAGGCTCTCA 1137  
QY 1142 GAGCAGAGAACTTGAAGAACAGATGCGGAGCTGCGCTGATCCCGCATGCTGTAC 1201  
DB 1138 GAGCAGAGAACTTGAAGAACAGATGCGGAGCTGCGCTGATCCCGCATGCTGTAC 1197  
QY 1202 GAGCTGACAGCAGCGCATCAAGTTTCATCAATGAACGGGCTTATGCGGACCCCATG 1261  
DB 1198 GAGCTGACAGCAGCGCATCAAGTTTCATCAATGAACGGGCTTATGCGGACCCCATG 1257  
QY 1262 AAGGTGTCAAAAGACCGGACAGTCAATGAACATGTGAGTGAAGCAGGAAGAGACCTTC 1321  
DB 1258 AAGGTGTCAAAAGACCGGACAGTCAATGAACATGTGAGTGAAGCAGGAAGAGACCTTC 1317  
QY 1322 CGGAGAGATTCATGACACATCCCAAGAACTTTGGCTGATCGCATCTTCTGGAGAGG 1381  
DB 1318 CGGAGAGATTCATGACACATCCCAAGAACTTTGGCTGATCGCATCTTCTGGAGAGG 1377  
QY 1382 AAGACATGTGCTGAGTGGTCTCTATTATTAATGACTGAATAAGAAATGAGAACTATAG 1441



[illegible]

D	247	GTGGTCCCCAAGAGAGAAAGAGAGAGACCCGACAGAGGCCCCCACTGTGAAGAGAGCG	2466
Q	2522	GAGAGACAGAAAGCCCCCGCGCTTGAGAGACTTGCGACTGAGACAAGAGAAAGCCGAGAG	2581
D	2467	GAGGAGCGAAAGCCCCCGCGCTGAGGAGCTGGCGACTGGACAAGAGAAAGCCGAGAG	2526
Q	2582	CCCGTCAAAGACCGAGTGCACGGAAGAAAGCCGAGAGAGGGGCGCGCAAGGGCAAGAACG	2641
D	2527	CCCGTCAAAGACCGAGTGCACGGAAGAAAGCCGAGAGAGGGGCGCGCTMAAGGCAAGAACG	2586
Q	2642	GAGGCGCGTGAAGGCCACGCGCGAGAGGGGCGCTCAAGGCGAGAAAGAAAGAGAGCGAGC	2701
D	2587	GAGGCGCGTGAAGGCCACGCGCGAGAGGGGCGCTCAAGGCGAGAAAGAAAGAGAGCGAGC	2646
Q	2702	GCGAGGGGCACACTGCGCAGAGCTCGGCGCGCCCCCAGAGCAGCGACTCCAGTGAAC	2761
D	2647	GCGAGGGGCACCAAGCGCAGAGCTCGGCGCGCCCCCAGAGCAGCGACTCCAGTGAAC	2706
Q	2762	TGCAGTGCAGACGAGGTGTGATGAGGCGCGAGGGGCGCGCAAGAACCGGCTGTCTCCCA	2821
D	2707	TGCAGTGCAGACGAGGTGTGATGAGGCGCGAGGGGCGCGCAAGAACCGGCTGTCTCCCA	2766
Q	2822	AGGCGCGAGCCCTCCACCCCGCAGCTGCGAGCCCGGGGCAATGCTCAAGCCCAAGGCA	2881
D	2767	AGGCGCGAGCCCTCCACCCCGCAGCTGCGAGCCCGGGGCAATGCTCAAGCCCAAGGCA	2826
Q	2882	CTGGAACCTGAAGCACTGAAGCAAGCAGCGGCTGCATCCCGCCCATTCAGGTCAACAA	2941
D	2827	CTGGAACCTGAAGCACTGAAGCAAGCAGCGGCTGCATCCCGCCCATTCAGGTCAACAA	2886
Q	2942	GTCCATGAGCCCCCGGAGAGACGCAAGCTCCACCAAGCAGTCCCCCAGCCCCACG	3001
D	2887	GTCCATGAGCCCCCGGAGAGACGCAAGCTCCACCAAGCAGTCCCCCAGCCCCACG	2946
Q	3002	CCACCGCAAAACCTGCAAGCGGAGAGAGAGGAGCGCCCTAGAGAGCTTGGAGAGGCCCG	3061
D	2947	CCACCGCAAAACCTGCAAGCGGAGAGAGAGGAGCGCCCTTAGAGAGCTTGGAGAGGCCCG	3006
Q	3062	GGCAAGAGCAAGAACCCGCGCACCCCGCGCAGCAAGAG-----3100	
D	3007	GGCAAGAGCAAGAACCCGCGCACCCCGCGCAGCAAGAGAGGAGAGGAGGCTGTCTTC	3066
Q	3101	---GCTTTCGAGCGGAGGCCCAGAGCTGCTGGGAGCCCCCTTGTGTGACTTCCGCG	3157
D	3067	CCAGCTTTCGAGCGGAGGCCCAGAGACTGCTGGGAGCCCCCTTGTGTGACTTCCGCG	3126
Q	3158	CTGCGCTTTCCTGCGGCCCGCGCGCTGAGAGTATCAAGGCTTCCCGCATGCCCCGAGACCC	3217
D	3127	CTGCGCTTTCCTGCGGCCCGCGCGCTGAGAGTATCAAGGCTTCCCGCATGCCCCGAGACCC	3186
Q	3218	TCAAGCTTTCCTGAGGCTTCAAGCTGAGTCAAGGCTTCCCGCATGCCCCGAGACCC	3277
D	3187	TCAAGCTTTCCTGAGGCTTCAAGCTGAGTCAAGGCTTCCCGCATGCCCCGAGACCC	3246
Q	3278	CGGCGCGTCTGCGCGCGCCACCCACCATCTCCAAACCGGCTTCCCATCTCTGCG	3337
D	3247	CGGCGCGTCTGCGCGCGCCACCCACCATCTCCAAACCGGCTTCCCATCTCTGCG	3306
Q	3338	AAGCAACCCAGAGCTTCTGAGAGGCAAAATGAGTGCATCTTCCAAAGAAATGTCCGTCA	3397
D	3307	AAGCAACCCAGAGCTTCTGAGAGGCAAAATGAGTGCATCTTCCAAAGAAATGTCCGTCA	3366
Q	3398	CTTCAAGTCCCGTACTAGAGCATGCAAGGCGCGGTTGGGCGCTGTCACATGGGCTG	3457
D	3367	CTTCAAGTCCCGTACTAGAGCATGCAAGGCGCGGTTGGGCGCTGTCACATGGGCTG	3426
Q	3458	CCCCTGCGCATGAGCCCAAAAGCTGCGACCTTTCAGCGAGTGAAGCAGAGCAGCTG	3517
D	3427	CCCCTGCGCATGAGCCCAAAAGCTGCGACCTTTCAGCGAGTGAAGCAGAGCAGCTG	3486
Q	3518	TCCCCAAGGGGCCAGGCTGGGCGCACCGAGAGAGCTGGGGGTGCTCCACAGCCCAAGAGCG	3577
D	3487	TCCCCAAGGGGCCAGGCTGGGCGCACCGAGAGAGCTGGGGGTGCTCCACAGCCCAAGAGCG	3546

Qy 3578 TCCGTGCTGAGAGGAGCACTCTGGGCTCAgTTCGGGGGAGAGCATCAcCAAAAGCAATT 3637  
Db 3547 TCCGTGCTGAGAGGAGCAgCTCTGGGCTCAgTTCGGGGGAGAGCATCAcCAAAAGCAATT 3606  
Qy 3638 CCCAGCAcACGGGTCCTCTCGAGcAGCCGCATCACTATCCGGGCTCCATCAcCCACGGC 3697  
Db 3607 CCCAGCAcACGGGTCCTCTCGAGcAGCCGCATCACTATCCGGGCTCCATCAcCCACGGC 3666  
Qy 3698 ACGCAGCTGAAGCTCTGTATCAAGGGCAcCATCAcCAGGATCATCGGAGGAGCAgCCCG 3757  
Db 3667 ACGCAGCTGAAGCTCTGTATCAAGGGCAcCATCAcCAGGATCATCGGAGGAGCAgCCCG 3726  
Qy 3758 AGTCGCTTGAACCGCGGCGCGGAGAGCAgCCTGCCAAAGGGCAcGTCACTTACGAAGC 3817  
Db 3727 AGTCGCTTGAACCGCGGCGCGGAGAGCAgCCTGCCAAAGGGCAcGTCACTTACGAAGC 3786  
Qy 3818 AAGAAAGGGCAcAGTCTTGTCTTATGAGGGTGAGATGTCTGTGACCCAGTCTTCAAGAG 3877  
Db 3787 AAGAAAGGGCAcAGTCTTGTCTTATGAGGGTGAGATGTCTGTGACCCAGTCTTCAAGAG 3846  
Qy 3878 GACGCGAGAAAGAGCTCAgGAGCCGCCCATGAGAGCGCGCGCCCAAGGGCAcCTATGAC 3937  
Db 3847 GACGCGAGAAAGAGCTCAgGAGCCGCCCATGAGAGCGCGCGCCCAAGGGCAcCTATGAC 3906  
Qy 3938 ATGATGAGAGGGCGCGTGAGGAGCAgTCTCTCAgCCAGCATCGAAAGTCTATGAGC 3997  
Db 3907 ATGATGAGAGGGCGCGTGAGGAGCAgTCTCTCAgCCAGCATCGAAAGTCTATGAGC 3966  
Qy 3998 CGTGCATTCGCCCGCGAGAGCAcAGCCGCCCAcCTTCAAAAGAGAGCAcCATTCGGC 4057  
Db 3967 CGTGCATTCGCCCGCGAGAGCAcAGCCGCCCAcCTTCAAAAGAGAGCAcCATTCGGC 4026  
Qy 4058 GGGTCATCACTCAcAAGAGGATCCCTCGATCTTACGTGAGAGCAcAGAGGACTTACCTGAGT 4117  
Db 4027 GGGTCATCACTCAcAAGAGGATCCCTCGATCTTACGTGAGAGCAcAGAGGACTTACCTGAGT 4086  
Qy 4118 CGGAGAGCCAAAGCTCTTAAAGCGGAGAGGCAcGCTTCGCGCCCAcCGGCTCAcGAGAGC 4177  
Db 4087 CGGAGAGCCAAAGCTCTTAAAGCGGAGAGGCAcGCTTCGCGCCCAcCGGCTCAcGAGAGC 4146  
Qy 4178 CTGACCGAGGCTTAAAGAGCGAGGCTCTGGGCCCCCTGAAAGCTGAAGCGCGCCATGAG 4237  
Db 4147 CTGACCGAGGCTTAAAGAGCGAGGCTCTGGGCCCCCTGAAAGCTGAAGCGCGCCATGAG 4206  
Qy 4238 GGCCTGTGTGGCAcCGGTGAAGAGGCGGAGCGCTCATCATGAGATCCCGGCGAGAG 4297  
Db 4207 GGCCTGTGTGGCAcCGGTGAAGAGGCGGAGCGCTCATCATGAGATCCCGGCGAGAG 4266  
Qy 4298 CTGCGGCAcAGCGCCGAGCTGCGCTTGCGCCCGCGGCGCTCAAGAGAGCTCATCAcG 4357  
Db 4267 CTGCGGCAcAGCGCCGAGCTGCGCTTGCGCCCGCGGCGCTCAAGAGAGCTCATCAcG 4326  
Qy 4358 CAGGGCAcCCCGCTCAAGTACGAcACCGGCGGTCAcCATGAGTCCAAAAAGCAcGAC 4417  
Db 4327 CAGGGCAcCCCGCTCAAGTACGAcACCGGCGGTCAcCATGAGTCCAAAAAGCAcGAC 4386  
Qy 4418 GTAAGCTCCCTCATGTGAGAGCCCGGCGGAGAGTTCcCAcCCGTCAGCCCGCTGAGTGG 4477  
Db 4387 GTAAGCTCCCTCATGTGAGAGCCCGGCGGAGAGTTCcCAcCCGTCAGCCCGCTGAGTGG 4446  
Qy 4478 ATGCGCGAcCGCCCGGCAcTGAGAcGTGTGCTTCAcGAGAGAGGCTGAAAGCGGCAc 4537  
Db 4447 ATGCGCGAcCGCCCGGCAcTGAGAcGTGTGCTTCAcGAGAGAGGCTGAAAGCGGCAc 4506  
Qy 4538 GGGAGCGCGCAcGAGCTCGGGGGGCTTCATTTGGCGCGGCGCGGCTCATTTGTGCTGAG 4597  
Db 4507 GGGAGCGCGCAcGAGCTCGGGGGGCTTCATTTGGCGCGGCGCGGCTCATTTGTGCTGAG 4566  
Qy 4598 CTGGGTAAAGCGCGGCAcGAGCCCGCTGAcCTTATGAGAcCAcAGGGGCAcCCCTTGGCGG 4657  
Db 4567 CTGGGTAAAGCGCGGCAcGAGCCCGCTGAcCTTATGAGAcCAcAGGGGCAcCCCTTGGCGG 4626

Qy 4658 CACCTCCACAGAGTTCGCCCGTGAcCATGCGGAGAGCCCAcCGCGGCTCGAGAGAGGC 4717  
Db 4627 CACCTCCACAGAGTTCGCCCGTGAcCATGCGGAGAGCCCAcCGCGGCTCGAGAGAGGC 4686  
Qy 4718 AGCCTTGTGTCAGAGAAAGCATCCAGAGACCGAAAGCTGAAGTGAAGCTCGTGAATC 4777  
Db 4687 AGCCTTGTGTCAGAGAAAGCATCCAGAGACCGAAAGCTGAAGTGAAGCTCGTGAATC 4746  
Qy 4778 GCCAAGTCCCGCAcAGAGACCGTGCcCGAGAGCAcCCCAcCCCATCTCGCCTATGAG 4837  
Db 4747 GCCAAGTCCCGCAcAGAGACCGTGCcCGAGAGCAcCCCAcCCCATCTCGCCTATGAG 4806  
Qy 4838 CACTGTCTTGGGGGTGAAGTGAAGCTGTATGAGAGCAcATCCCTGGGCTTC 4897  
Db 4807 CACTGTCTTGGGGGTGAAGTGAAGCTGTATGAGAGCAcATCCCTGGGCTTC 4866  
Qy 4898 GACCCCAcCTCATTCcCGGGGCAcTCCCTGAGAGGAGCGGCTGACCTACCTGCGC 4957  
Db 4867 GACCCCAcCTCATTCcCGGGGCAcTCCCTGAGAGGAGCGGCTGACCTACCTGCGC 4926  
Qy 4958 CGACACTGGCCCCCAcCCCAcCTACCGGACCTGTACCCCAcCTCATTCcCGGCGC 5017  
Db 4927 CGACACTGGCCCCCAcCCCAcCTACCGGACCTGTACCCCAcCTCATTCcCGGCGC 4986  
Qy 5018 TACCCCGAcAGCGCGGCTGAGAAcCGGCAcCATCATCATATGACTCATCAcCTCG 5077  
Db 4987 TACCCCGAcAGCGCGGCTGAGAAcCGGCAcCATCATCATATGACTCATCAcCTCG 5046  
Qy 5078 CAGAGATGCAcCAcAAGCGGCAcCGGCAcTGAAGCGGAGCTGATGCTGAGGGGC 5137  
Db 5047 CAGAGATGCAcCAcAAGCGGCAcCGGCAcTGAAGCGGAGCTGATGCTGAGGGGC 5106  
Qy 5138 CTCTCGCCCGAGTCTCTGCTGCACTCACTTACGCTGCGGAGTCCCCAGGAGCATATC 5197  
Db 5107 CTCTCGCCCGAGTCTCTGCTGCACTCACTTACGCTGCGGAGTCCCCAGGAGCATATC 5166  
Qy 5198 GACTGTTCcCAAGTGCACAcCTGTGTGCTCGTCCCGGACCAcAGGACCCCAcGC 5257  
Db 5167 GACTGTTCcCAAGTGCACAcCTGTGTGCTCGTCCCGGACCAcAGGACCCCAcGC 5226  
Qy 5258 ACCGCAcTGAAGCGGCTTGCCTTGCCTCCCAcCGGCGCCAGCCCTTGAAGAGCGGCAc 5317  
Db 5227 ACCGCAcTGAAGCGGCTTGCCTTGCCTCCCAcCGGCGCCAGCCCTTGAAGAGCGGCAc 5286  
Qy 5318 AGCAGCTCCCAcCTCTCCCGAGAGTCCAAcCAcCTTGAcAAAAcCAcCAcCAcGTC 5377  
Db 5287 AGCAGCTCCCAcCTCTCCCGAGAGTCCAAcCAcCTTGAcAAAAcCAcCAcCAcGTC 5346  
Qy 5378 TCGTCCAGCGGAGCGAGAcCCGGGATGAGAGCGGAGCCGGGATCGGAGCGGAAAAAG 5437  
Db 5347 TCGTCCAGCGGAGCGAGAcCCGGGATGAGAGCGGAGCCGGGATCGGAGCGGAAAAAG 5406  
Qy 5438 TCCATCCcCAGTCCACAGAcCGGTGAGAcCGCAcCATCTGAGAcCTGTGACAGAG 5497  
Db 5407 TCCATCCcCAGTCCACAGAcCGGTGAGAcCGCAcCATCTGAGAcCTGTGACAGAG 5466  
Qy 5498 CAGAGCAcCGGCAcGAGCGGAGCAcCGGCGGAGTGGGGGCAcGAGCGGCCCCCGC 5557  
Db 5467 CAGAGCAcCGGCAcGAGCGGAGCAcCGGCGGAGTGGGGGCAcGAGCGGCCCCCGC 5526  
Qy 5558 TCCCAcCTCCATGcCCACAGAcCTGCGCATCTTCCTTGAGAcCCAGATGCTTCAG 5617  
Db 5527 TCCCAcCTCCATGcCCACAGAcCTGCGCATCTTCCTTGAGAcCCAGATGCTTCAG 5586  
Qy 5618 CAGAGAcCAcTGTGCTTCAcAAGCAcAGGAGTGAAGGTTATCATCAcCGGTGTGAGGCC 5677  
Db 5587 CAGAGAcCAcTGTGCTTCAcAAGCAcAGGAGTGAAGGTTATCATCAcCGGTGTGAGGCC 5646  
Qy 5678 AGCAAGCCCAcGAGTCTGAGTCCAcCTCAcCTTCAcCCGTTGCGCAGCTCCACA 5737  
Db 5647 AGCAAGCCCAcGAGTCTGAGTCCAcCTCAcCTTCAcCCGTTGCGCAGCTCCACA 5706  
Qy 5738 TTCCAcCTGcCAcCCAcCTGcCCAcCTGGGCGGAcCCCTGATGAGGAGTCAcCTTACCTC 5797

```

Db      5707 TTCCCACTGCAACCCACTGACCCTGGAGGCACTCGATGGGTCTACCTTACCTTC 5766
Qy      5798 ATGGAGCCCGCTCTTGCTGCTCCCAAGAGGCCCCCGGGTGGCCCGGCAAGCGGCCCCCA 5857
Db      5767 ATGGAGCCCGCTCTTGCTGCTCCCAAGAGGCCCCCGGGTGGCCCGGCAAGCGGCCCCCA 5826
Qy      5858 GCAGACACCGGCAATGCTTCTCTGCGCAAGCCCCGAGCCCGCTCGGGGTGAGAGCCGCG 5917
Db      5827 GCGAGACCGGGCAATGCTTCTCTGCGCAAGCCCCGAGCCCGCTCGGGGTGAGAGCCGCG 5886
Qy      5918 TCCCTCCCGAGAAAGGCTTCGAGCCCGGCGCTGATGCTCTCTGCGCAAGCG 5977
Db      5887 TCTCTCCCGAGAAAGGCTTCGAGCCCGGCGCTGATGCTCTCTGCGCAAGCG 5946
Qy      5978 ACCATGCCCCGACCCCTGCGAAAGAACTTCGACCTTCACACGCGAGCCCGGAGCCGCG 6037
Db      5947 ACCATGCCCCGACCCCTGCGAAAGAACTTCGACCTTCACACGCGAGCCCGGAGCCGCG 6006
Qy      6038 GCGCCACCTGCGCTCGGCTTCGAGCCCGGCAACCGGAAAGAACTCAAGTAAACCTTTTCC 6097
Db      6007 GCGCCACCTGCGCTTCGAGCTTCGAGCCCGGCAACCGGAAAGAACTCAAGTAAACCTTTTCC 6066
Qy      6098 ATCCAGGAATGGAATCTCCGTTCTCTGGGTTCACACGCGAGCACTTACAGCCCGGAGG 6157
Db      6067 ATCCAGGAATGGAATCTCCGTTCTCTGGGTTCACACGCGAGCACTTACAGCCCGGAGG 6126
Qy      6158 GTGGAGCCCGCTGAGCCCTGTGAGCTACCCAGTTCGACCCAGCAAGGGGCTCCCCAG 6217
Db      6127 GTGGAGCCCGCTGAGCCCTGTGAGCTACCCAGTTCGACCCAGCAAGGGGCTCCCCAG 6186
Qy      6218 CACTGGAAGAGTGCAGCAAGAGCCACTGAGGGGAGCTGGCGGCCCAAGAGCCAGG 6277
Db      6187 CACTGGAAGAGTGCAGCAAGAGCCACTGAGGGGAGCTGGCGGCCCAAGAGCCAGG 6246
Qy      6278 CCGGTGAAGCTTGGCGGGAGGCGGCCCACTTCCCAACCTGGCGCGCTGCTGAGAGC 6337
Db      6247 CCGGTGAAGCTTGGCGGGAGGCGGCCCACTTCCCAACCTGGCGCGCTGCTGAGAGC 6306
Qy      6338 CAGCCCTCGTCCAGCCCGCTGCTCCAGACCGGCCCGAGGGGTCAAGAGTCAACAGCGGGT 6397
Db      6307 CAGCCCTCGTCCAGCCCGCTGCTCCAGACCGGCCCGAGGGGTCAAGAGTCAACAGCGGGT 6366
Qy      6398 GTCAACCTGGGCCAGGACATCATGAGGTTCATCAACAGGACTACCCGGGCAACCCCA 6457
Db      6367 GTCAACCTGGGCCAGGACATCATGAGGTTCATCAACAGGACTACCCGGGCAACCCCA 6426
Qy      6458 CAGCAGCTCAGGCGCACCCCTGGCGGCCCTCTACTCTCTTCCCTGGGGCAGCTGCCCC 6517
Db      6427 CAGCAGCTCAGGCGCACCCCTGGCGGCCCTCTACTCTCTTCCCTGGGGCAGCTGCCCC 6486
Qy      6518 GTCTGTGACCTTCGCGCGGCCCAACAGTGACTTCACTTCCCGGCCCGAGCAATGGTGCC 6577
Db      6487 GTCTGTGACCTTCGCGCGGCCCAACAGTGACTTCACTTCCCGGCCCGAGCAATGGTGCC 6546
Qy      6578 CCGGCCCGGTGCTCCCGGCAAGGGGGGAGAGAGGTTCAGAGGCAACAAAGAG 6637
Db      6547 CCGGCCCGGTGCTCCCGGCAAGGGGGGAGAGAGGTTCAGAGGCAACAAAGAG 6606
Qy      6638 TCGGTCTTGGGTGTGTGTGAGGACGGTATTGAACTGTGTCCCAACCGAGGGGCAATGACG 6697
Db      6607 TCGGTCTTGGGTGTGTGTGAGGACGGTATTGAACTGTGTCCCAACCGAGGGGCAATGACG 6666
Qy      6698 GAGCCAGGGCACTCCCGAGGTGTGTATCCGCTGTCTTACCGGGATGGGAAACAGACG 6757
Db      6667 GAGCCAGGGCACTCCCGAGGTGTGTATCCGCTGTCTTACCGGGATGGGAAACAGACG 6726
Qy      6758 GAGCCAGGAGATGGGCTCAAGTCTCAGGCAACAGAGCCAGCGGCAAGCTTCTTC 6817
Db      6727 GAGCCAGGAGATGGGCTCAAGTCTCAGGCAACAGAGCCAGCGGCAAGCTTCTTC 6786
Qy      6818 AGCAAGCTGACGAGCAACTCCGCAATGTCAAGTCCAAAGAGCAAGATCAACAAG 6877

```

```

Db      6877 AGCAAGCTGACGAGCAACTCCGCAATGTCAAGTCCAAAGAGCAAGATCAACAAG 6846
Qy      6878 AAGCTGAACCCCAACACCGGAATGAGCCTGAATCAATATACAGGCTGGGAGGAG 6937
Db      6847 AAGCTGAACCCCAACACCGGAATGAGCCTGAATCAATATACAGGCTGGGAGGAG 6906
Qy      6938 ATCTTCAATATGCTCCGCTCATCAACCGGAACAGGCTTATGACCTATGAAAGCCAGCGGT 6997
Db      6907 ATCTTCAATATGCTCCGCTCATCAACCGGAACAGGCTTATGACCTATGAAAGCCAGCGGT 6966
Qy      6998 CAGGAACATGCAACCAACATGGGGCTGAGAGCCATATTAAGAGCACTATGGGT 7057
Db      6967 CAGGAACATGCAACCAACATGGGGCTGAGAGCCATATTAAGAGCACTATGGGT 7026
Qy      7058 AATATGACAGTGGGAAGAGTCCCGGCGCTCAGGCGCAAGGCTTTTAACTCTGAAT 7117
Db      7027 AATATGACAGTGGGAAGAGTCCCGGCGCTCAGGCGCAAGGCTTTTAACTCTGAAT 7086
Qy      7118 GCGAGTGCAGGCTGCGCGCTGTATGACCATTAACCGCTGTGACGAGCGAGTGCAC 7177
Db      7087 GCGAGTGCAGGCTGCGCGCTGTATGACCATTAACCGCTGTGACGAGCGAGTGCAC 7146
Qy      7178 AACTCACTCGCCAGTGGCGGCGGAAAGCCAAAGTCTTGGCAGACCCAGCGCGCA 7237
Db      7147 AACTCACTCGCCAGTGGCGGCGGAAAGCCAAAGTCTTGGCAGACCCAGCGCGCA 7206
Qy      7238 AAAGCCAAATCCCGGCGCGCGGCGCTGAGATCTGGGGAACCGGCCCTGTCTCTCA 7297
Db      7207 AAAGCCAAATCCCGGCGCGCGGCGCTGAGATCTGGGGAACCGGCCCTGTCTCTCA 7266
Qy      7298 GTGCACTCGAAGGAGACTGCAACCGCGGAGCGGCTCAACACCGGCTGTGGAGAGC 7357
Db      7267 GTGCACTCGAAGGAGACTGCAACCGCGGAGCGGCTCAACACCGGCTGTGGAGAGC 7326
Qy      7358 AGGCTCTGTCGCAAGTTCAGGCTTATCCCTTCAACACCCCTGTATCATGGGCTGCAG 7417
Db      7327 AGGCTCTGTCGCAAGTTCAGGCTTATCCCTTCAACACCCCTGTATCATGGGCTGCAG 7386
Qy      7418 GGGGTGTCAATGGCTTTCCCAACCCCAACGGGCTTCCCGCGGGGAGCGGGCCCTCGCT 7477
Db      7387 GGGGTGTCAATGGCTTTCCCAACCCCAACGGGCTTCCCGCGGGGAGCGGGCCCTCGCT 7446
Qy      7478 GGGCCCAACAGGCTGGGACGAGAGCCCAAGCACTGTGTGCTGCAATACAGACA 7537
Db      7447 GGGCCCAACAGGCTGGGACGAGAGCCCAAGCACTGTGTGCTGCAATACAGACA 7506
Qy      7538 CTCTCGACAGCGAG 7552
Db      7507 CTCTCGACAGCGAG 7521

```

Search completed: March 12, 2006, 06:25:27  
 Job time : 3076 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2006, 12:21:21 ; Search time 2098 Seconds  
(without alignments)  
9426.045 Million cell updates/sec

Title: US-09-522-753-4  
Perfect score: 8564.8  
Sequence: 1 catgctgggcccacacagc.....caaaaaaaaaaaaaaaaaa 8571

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7673375 seqs, 115364844 residues

Total number of hits satisfying chosen parameters: 15346750

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA New:  
1: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	630	7.4	7940	US-11-245-147-191	Sequence 191, App
2	190.6	2.2	559	US-11-136-527-1693	Sequence 1693, App
3	190.6	2.2	559	US-11-136-527-5789	Sequence 5789, App
4	186.2	2.2	609	US-09-925-065A-841377	Sequence 841377, App
5	156.8	1.8	421	US-09-925-065A-516005	Sequence 516005, App
6	126.2	1.5	600	US-10-750-185-1127	Sequence 1127, App
7	126.2	1.5	600	US-10-750-623-1127	Sequence 1127, App
8	115	1.3	30191	US-10-330-773-631	Sequence 631, App
9	113.4	1.3	72352	US-10-330-773-43	Sequence 43, App
10	112.4	1.3	302	US-09-925-065A-516003	Sequence 516003, App
11	98.6	1.2	153376	US-11-121-086-5	Sequence 5, App
12	84	1.0	1061	US-09-925-065A-720226	Sequence 720226, App
13	83.6	1.0	409	US-09-925-065A-16738	Sequence 16738, App
14	82.4	1.0	2093	US-11-072-512-1666	Sequence 1666, App
15	81.2	0.9	1400	US-11-136-527-3376	Sequence 3376, App
16	81.2	0.9	2280	US-11-136-527-3376	Sequence 3376, App
17	79	0.9	204603	US-10-330-773-325	Sequence 325, App
18	76	0.9	1728	US-10-750-185-42639	Sequence 42639, App
19	76	0.9	1728	US-10-750-623-42639	Sequence 42639, App
20	75.8	0.9	118864	US-10-330-773-826	Sequence 826, App

21	74.8	0.9	2539	US-11-136-527-215	Sequence 215, App
22	73.4	0.9	26214	US-10-330-773-222	Sequence 222, App
23	73	0.9	1027	US-11-136-527-94	Sequence 94, App
24	71.8	0.8	2183	US-11-072-512-1064	Sequence 1064, App
25	71.6	0.8	2479	US-11-136-527-3303	Sequence 3303, App
26	71.6	0.8	9474	US-11-052-554A-526	Sequence 526, App
27	71.6	0.8	194186	US-10-330-773-395	Sequence 395, App
28	70.8	0.8	7474	US-11-069-834-49	Sequence 49, App
29	70.6	0.8	600	US-11-136-527-7399	Sequence 7399, App
30	70.2	0.8	610	US-09-925-065A-780159	Sequence 780159, App
31	70.2	0.8	610	US-09-925-065A-780160	Sequence 780160, App
32	69.2	0.8	365	US-11-043-752-3906	Sequence 3906, App
33	69.2	0.8	365	US-11-043-752-3909	Sequence 3909, App
34	68.6	0.8	4146	US-11-052-554A-522	Sequence 522, App
35	68.2	0.8	657	US-09-925-065A-952675	Sequence 952675, App
36	67.8	0.8	1619	US-10-821-234-10	Sequence 10, App
37	67.8	0.8	2772	US-11-052-554A-531	Sequence 531, App
38	67.6	0.8	26667	US-10-995-561-13375	Sequence 13375, App
39	67.4	0.8	244	US-11-043-752-3912	Sequence 3912, App
40	66.4	0.8	207908	US-11-112-908-19	Sequence 19, App
41	66.4	0.8	212805	US-11-112-908-19	Sequence 19, App
42	66.2	0.8	5706	US-11-052-554A-519	Sequence 519, App
43	66	0.8	649	US-09-925-065A-720225	Sequence 720225, App
44	66	0.8	38920	US-10-330-773-228	Sequence 228, App
45	65.8	0.8	5706	US-11-052-554A-519	Sequence 519, App

ALIGNMENTS

RESULT 1  
US-11-245-147-191  
Sequence 191, Application US/11245147  
Publicat ion No. US20060030541A1  
GENERAL INFORMATION:  
APPLICANT: GARCIA, TERESA  
APPLICANT: ROMAN ROMAN, SERGIO  
APPLICANT: BARON, ROLAND  
APPLICANT: CALL, KATHERINE  
APPLICANT: THEILHABER, JONCHIM  
APPLICANT: CONNOLLY, TIMOTHY  
APPLICANT: JACKSON, AMANDA  
APPLICANT: BUSHNELL, STEVEN  
APPLICANT: RAMADI, GEORGES  
TITLE OF INVENTION: GENES INVOLVED IN OSTEOGENESIS, AND METHODS OF USE  
FILE REFERENCE: 37991-0023  
CURRENT FILING DATE: 2005-10-07  
PRIORITY FILING DATE: 2005-10-07  
PRIORITY FILING DATE: 2002-04-05  
PRIORITY FILING DATE: 2002-04-05  
PRIORITY FILING DATE: 2001-04-05  
NUMBER OF SEQ ID NOS: 246  
SOFTWARE: Patent In Ver. 3.2  
SEQ ID NO 191  
LENGTH: 7940  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Homo sapiens nuclear receptor co-repressor 1  
OTHER INFORMATION: (NCOR1), mRNA  
US-11-245-147-191

Query Match 7.4%; Score 630; DB 9; Length 7940;  
Best Local Similarity 65.1%; Pred. No. 4.8e-115;  
Matches 1015; Conservative 0; Mismatches 495; Indels 48; Gaps 4;

QY 491 CCGCAACGGCTGTCCAGAGAGAGCTGATCCAGACATGAGCCGGGTGAGACGAGAGATC 550  
DB 757 CTTGAAACCTCTCAAGAGAGAGATTAATACAGAGTATGAGATCGTGTGATGAGAAATT 816  
QY 551 ACCATGATGAGACAGATCTCTAAGCTGAAGAGAGACAGCAACAGCTGAGAGAGAG 610

Db	817	GCAGAAAGTGAACACGACGATCCTTTAACTGAGAAAAAGAAACAACACGCTTGAAAGAGG	876
Qy	611	GCTGCCAAGCCGCCCGAGCCTTGAGAAAGCCCGTGTCACGCGCGCCATCGAGTTCGAAGAC	670
Db	877	GCAGCTAAACCTCCTGAGCCTTGAGAAAGCCCGTGTCCTCCCTCTCTGTGGAGCAAGAAC	936
Qy	671	CGCAGCCCTGTGTGCANATCATCTACGACGAGAACCCGGAAGAAAGCTGMAAGTGCACATCGG	730
Db	937	CGCAGTATTTGTCCAAATTTATTTATATGAGTAATCGAAAAAGCAGAAACACTCATTA	996
Qy	731	ATTCTGAAAGCCTTGAGGCCCGCAGGTGAGCTGCGCTGTACAAACAGGCCCTCCGACACC	790
Db	997	ATTTTGAAGGCTTGGGCCCAAAAGTTGACCTGCCACTGTATTAACAGGCATCAAGTACC	1055
Qy	791	CGCGATATCATGAGAACATCAAAATTAACCGAGCGATGCGGAGAGAGCTAACTTGTAC	850
Db	1057	AAGGTGTACCATGAGAACATCAAGCAAAACGAGTGATGAGGAAAAAACTCATTTTATTT	1111
Qy	851	TTTCAAGAGAGAAATCACGCTTCGAGAAACATGGAAGCAGAAGTTCTGCGAGGCTATGAC	910
Db	1117	TTTAAAAAGAAATCATGCAAGAAAACAAAGGAAACAAAAATCTGCCAGGTTATGAT	1177
Qy	911	CAGCTCATGAGGCGCTTGAAAAAAAAGGTGAGCCGATCGAAAAAAACCCGCGCGCGCG	970
Db	1177	CAGCTCATGAGGCGATGAGGAAAAAAGTGACAGATATGAAAAATATCTCTCGAGGAAA	1233
Qy	971	GCCAAAGAGACCAAGGTGCGCGAGTACTACGAAAAAGCAAGTTCTCTGAGATCCGCAAGCAG	1030
Db	1237	GCTAAAGAAACAAAACAAGGAAATCATATGAAAAAGCAGTTTCCAGAAATTCGAAAAACA	1296
Qy	1031	CGCGAGCTGCAAGAGCCGATGCAGACGAGGCTGGGCGACGGGGCAGTGGCTGTTCATG	1091
Db	1297	AGAGAACGCAAGAAAGATTTTCAG---CGAGTTGGGCAAGAGGGGAGCTGTCTTTTCAGCG	1353
Qy	1091	TGCGCGCGCGCAGCGAGACGAGGCTGACAGATCATGATGGGCTCTCAAGACAGAG	1155
Db	1354	ACCATGCTTAGAGATGACAGATGAGATTTCTGAAATTTATGATGGGCTCTCTGAGCAGAG	1411
Qy	1151	AACCTGAGAAAGCAGATGCGCGAGCTGGCCGTGATCCGCGCCATGCTGTACGACGCTGAC	1210
Db	1414	AATATATGAAACAACAATGCGGAGCTCTCTGTGATTCACCTATGATGTTGATGTCAGAA	1477
Qy	1211	CAGCAGCGCATCAGTTTATCATCAACATGAACGGGCTTTATGGCCGACCCCATGAAAGTGTAC	1277
Db	1474	CAAGACACAGTCAAGTTTCATTAACTGATGGGCTTATGAGAGACCTATGAAAGTGTAT	1533
Qy	1271	AAAGACCGCCAGTGCATGAACATGTGGAGTGACGAGAGAGAGAACCTTCCGGAGAGAG	1333
Db	1534	AAAGATATGAGCGTTTATATGAATGTTTGGACTGACATGAAGAAAGAGATCTTTAAGACAG	1593
Qy	1331	TTTCATGCAAGATCCCAAGAACTTTGGCCGTGATCGATCATTTCTGAGAGAGAAAGCAGTG	1390
Db	1594	TTTATCCAGCATCCAAAAAATTGGACTTAATTCATCTTGGAGAGAAAGTGT	1655
Qy	1391	GCTGATGTCGCTCTTATTACTACCTGACTAAGAAATGAGAACTTATGAGCCTGTGTG	1455
Db	1654	CCTGATGTTGTTTGTATTACTAATTTAACAAAGAAAAATGAGAAATTAATAAGCCTCGTC	1711
Qy	1451	AGACGAGCTATCGGCGCGCCGCGCAAGAGCCAGAGGCAACAACAGCAGCAGCAGCAGCAG	1511
Db	1714	AGAGAGAAATTTATGGAAAAACGAGAGGCAAGAAACCGCA-----AATTGCTCGACCT	1765
Qy	1511	CAGCAGCAGCAGCAGCAGCAGCCCTGCGCCGACAGCAGCAGGAGAGAAAGATGAGAG	1577
Db	1766	CGCAAGAAAGAAAGTATGAAAGAAAGAAAGAGATTAAGCAGAAAAAACAGAAAAAAG	1822
Qy	1571	GAGAAAGAAAGAGCGGAGAGAGAGAGAGAAAGCCGAGGTGAGAAACGACAAAGAA	1633
Db	1826	AAAGAAAGAAA-----GAAAGTGAAGGAAAAAAGATGAAGAAAGAGATCTCAAAGAA	1878
Qy	1631	GACCTCTCTCAAGAGAGAGACAGACACCTTCAGGGGAGAGACAGACGACGAGAGAGGCT	1691
Db	1879	AATTCACAGAGAAAGAGACAGATATGATGTATCAGCAGAAAGAACTGAGAGAAAGAGACA	1933

QY	1691	GTGGCTTC	CGAAAGCG	CGAAATG	CGCAAC	CGGAGG	CGCGAT	CAACC	1750
Db	1939	GCCACACCC	CGGGGGCG	CGAAGAC	GTGCCAA	CAGTCA	GGGGCC	CGGTAA	GGGGCTCAC
QY	1751	CGCTCAAT	GGCTAATGA	-----	-----	-----	-----	GGCAACAGCGAG	1780
Db	1999	AGGTTCAT	GAAACG	AAAGCT	GTGAC	GTGCGC	CGGACG	CGCGAT	CGCTGAAAG
QY	1781	GAGGCGAT	CAACCCCG	CAGCAG	CGCGCG	CGAGCT	CTCATG	AGCTGAA	TGATGAGT
Db	2059	CCCCACCA	CACTTGTG	CCACCG	CCACGAA	CACTTCTT	TAAGAG	CGCTGTG	GAACCTCT
QY	1841	CGCTGACA	GAAAGAA	GAATGAAA	CAGCCAA	GAAGTCT	CTTGAA	CA	CGGCGCAAC
Db	2119	CGATGACA	GAGAAAG	AAATGAA	GTGCTAAA	AAAGCT	CTAGTA	GAAACAT	GTGTCTAAC
QY	1901	TGTGTGCG	CAATCG	CCCGAT	GTGTGG	CTCCAA	GA	CTGTGT	GGCAGTGA
Db	2179	TGGGAGCA	AAATGCT	TAATGTTG	GGAGAA	CGAAAGT	GAAGCT	CAATGTA	ATTAATCTTAT
QY	1961	TTCAACTA	CAAGAA	AGGCGA	GAACCT	CGATGAG	ATCTTG	TCAGCAG	CAACCTGAGA
Db	2239	TTTAACAT	CTATAA	AGGCGA	CAACAT	CTTGACA	CACTCT	TACAGCAG	CACTAACA
RESULT 2									
US-11-136-527-1693/c									
; Sequence 1693, Application US/11136527									
; Publication No. US20050287570b1									
; GENERAL INFORMATION:									
; APPLICANT: Wyeth									
; APPLICANT: Mounts, William M									
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes									
; FILE REFERENCE: 031896-041000 (AM101086)									
; CURRENT APPLICATION NUMBER: US/11/136,527									
; CURRENT FILING DATE: 2005-05-25									
; PRIOR APPLICATION NUMBER: US 60/574,294									
; PRIOR FILING DATE: 2005-05-26									
; NUMBER OF SEQ ID NOS: 362830									
; SOFTWARE: PatentIn version 3.2									
; SEQ ID NO 1693									
; LENGTH: 559									
; TYPE: DNA									
; ORGANISM: Rattus norvegicus									
; FEATURES:									
; NAME/KEY: misc_feature									
; LOCATION: (543)..(543)									
; OTHER INFORMATION: n is a, c, g, or t									
US-11-136-527-1693									
Query Match 2.2%; Score 190.6; DB 12; Length 559;									
Best Local Similarity 68.6%; Pred. No. 2.4e-28;									
Matches 395; Conservative 20; Mismatches 135; Indels 26; Gaps 11;									
QY	7997	CCGCTTACGCTCTGGA	CAGATGGA	CCCAAGC	-CCTGTCCAG	CCCCCGCAGTGCCTCTTC	8055		
Db	555	CCCGCTGGCGTNTG	GAACA	GACACGTT	CCCACTTAT	CCTGCCCATGCTGTATC	496		
QY	8056	CGGTGCCCA	GA	CTGCCCGCAG	CCA	CGAGATTGCTG	GAACCA	AGTCA	GGCCAGGTGGG
Db	495	CAGTCCCA	AAAGG	CCACCCAC	AGCCACAR	-----GAC	TGGAGCCCAT	CRGAC	CGAGTGGG
QY	8116	CGACA	AAAGG	CCAGATG	CGGCGCTTG	GGGGGA	AGGATGCT	CGAGAG	ACTGATGTTT
Db	440	TGCAC	-AAAGG	CGCGGTAT	GGCTT	-GGGMA	CGATGTTT	GCAGAA	CCGACCGTTC
QY	8176	TTTCA	CATGCTGTG	CGCGAG	CGGTGG	GAAGAA	AGCAGATG	TAAATGAT	GTGTGGTT
Db	383	ATTCCG	CGCGCTGTG	MTATG	GCAT	-GGCG	GAAGG	CAAGT	TTAAATG
QY	8236	TACAGG	GTATATTTT	TGATAC	CTTCA	TGAATTA	TTATCA	GAATGTTT	TACGAC



Db	326	WACAGGGATATTTTGTGTACTTCAATGAGTATTAATCAGTCGGTTCACACAAAGGAAGA	267
QY	8296	CTTACCCAGATATTACTGCTGTGCTTGTATCTCTGCTTACCGTTAAAGGCGGTG	8355
Db	266	CTACCC-GTGTACTGCTGCTGTGC-TTGTGTMTTGTCTTCTGTCCCAAGGCAATGTG	209
QY	8356	CAGGCGGACAGTCSGTGAACCCATCACTCGCAGAGCCAAAGGGGCGGGGACTGTCGTCA	8415
Db	208	CCGGCTCA--AGGCGGTGGCCCMCCATTCGACAGACTGGGGGGCAGGGGMMCCGGCA---	154
QY	8416	CGCCCGCTGTGTCCTCCCTCCCTCCCTTCTTGGGAGATGAAATTGAAAGCGTATTC	8475
Db	153	--CAGGGACTGTGCTCCCTCCCTCTCTTCTTGGGCAAAATGACATTGAATGCTATTC	96
QY	8476	GTGGCGGCAATTTGCGCAGGTGTGTGATTTCTGTCAATTAACAACGTGTTCTAATTA	8535
Db	95	GYGGCGGCTCTGTC-----AYGGCGGCAATTCGTGATTTACATGTTGTTCCAAATTA	41
QY	8536	AAAGCGAATTAATCTCCAAAAAAAAAAAAAAAAAAAAA	8571
Db	40	AAAGCAATTAATCTCCAAATTTAAAAAAAAAAAAA	5

```

RESULT 3
US-11-136-527-5789
; Sequence 5789, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounats, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (LM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5789
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (17)..(17)
; OTHER INFORMATION: n is a, c, g, or t
; US-11-136-527-5789

```

Query Match	2.28	Score 190.6	DB 12	Length 559
Best Local Similarity	68.68	Pred. No. 2.4e-26		
Matches 395	Conservative 20	Mismatches 135	Indels 26	Gaps 11
Qy	7997	CCGCTTACCGCTCTGCAGACAGATGAGCAGGAC - CCTGTCCAGCCCCCAGTGCCTCGTTC	8055	
Dy	5	CCCGCTGCGCGCTATGACAGACAGACGTTCCAGCTTATCTGCGCCATCGCTCTCATCG	64	
Qy	8056	CGGTGCCCAAGACTGCCCCCAGCCAAAGATTTGCTGAAAACCAAGTCAGGCCAGGTGGG	8115	
Dy	65	CAGTCCCCAAAGGCCACCCAGCCCAAR ----- GATGGGAGCCCATCTRACCAAGTGGG	119	
Qy	8116	CGAACAAGAGGCGCAGGTGCGGCTCTGGGGGGGAACGATGCTCCGAGACATGACTGTTTT	8175	
Dy	120	TGCA - AAAAGGCGCGGAGTATGGCT - GGGGMAAGATGTTTGCAGGAACCGACCGTTC	176	
Qy	8176	TTTCACACATCGTTTGGCGCAGCGGTGGGAAGGAAGCAAGTGTAAATGATGTGTGGTT	8235	
Dy	177	ATTCCCGGCGCTGTGAGTATGGCAAT - GCGGGGAAGGCAAGGTTKTAATG - GTGTGGCT	233	
Qy	8236	TACAGGGTATATTTTGTATACCTTCAATGATTAATTCAGATGTTTTTACGCAAGAAAGA	8295	
Dy	234	MACAGGGTATATTTTGTATACCTTCAATGATTAATTCAGTGGCTTCAACCAAGAAAGA	293	
Qy	8296	CTTACCCAGTATTACTGCTGCTGTGCTTTTGATCTTGCTTACCGTTCAAGAGCGGTGTG	8355	

[illegible]

```

RESULT 4
US-09-925-065A-841377/c
, Sequence 841377, Application US/09925065A
, Publication No. US20040181048A1
, GENERAL INFORMATION:
, APPLICANT: Mang, David G.
, TITLE OF INVENTION: Identification and Mapping of Single
, FILE REFERENCE: 108827.135
, CURRENT APPLICATION NUMBER: US/09/925,065A
, CURRENT FILING DATE: 2001-08-08
, PRIOR APPLICATION NUMBER: US 60/243,096
, PRIOR FILING DATE: 2000-10-24
, PRIOR APPLICATION NUMBER: US 60/252,147
, PRIOR FILING DATE: 2000-11-20
, PRIOR APPLICATION NUMBER: US 60/250,092
, PRIOR FILING DATE: 2000-11-30
, PRIOR APPLICATION NUMBER: US 60/261,766
, PRIOR FILING DATE: 2001-01-16
, PRIOR APPLICATION NUMBER: US 60/289,846
, PRIOR FILING DATE: 2001-05-09
, NUMBER OF SEQ ID NOS: 957086
, SOFTWARE: FastSeq for Windows Version 4.0
, SEQ ID NO 841377
, LENGTH: 609
, TYPE: DNA
, ORGANISM: Homo sapiens
US-09-925-065A-841377

```

Query Match	2.2% ;	Score 186.2 ;	DB 6 ;	Length 609 ;
Best Local Similarity	98.4% ;	Pred. No. 1.8e-27 ;		
Matches 186 ;	Conservative 0 ;	Mismatches 3 ;	Indels 0 ;	Gaps 0 ;
QY	7185	CCTCGCCAGGTGACGCGCGGAGAGCCCAAGTCTCTTGACAGACCAGCAGCGGAAAGCCA	7244	
Db	212	CCCTGTGAGGTGGCGGCGGAGAGGCCCAAGTCTCTTGACAGACCAGCAGCGGAAAGCCA	153	
QY	7245	AGTCCCGCGGCGCGCGGCGTGGCAATCTGGGGACCGGCCACCTCTGTCTCTCAATGACAT	7304	
Db	152	AGTCCCGCGGCGCGCGGCGTGGCAATCTGGGGACCGGCCACCTCTGTCTCTCAATGACAT	93	
QY	7305	CGAGAGGAGACTGCAACCGCGGACCGCGTCTCAACCAACCGCTGTGGAGGACAGAGCCCT	7364	
Db	92	CGAGAGGAGACTGCAACCGCGGACCGCGTCTCAACCAACCGCTGTGGAGGACAGAGCCCT	33	
QY	7365	CGTCCGCAAGT	7375	
Db	32	CGTCCGCAAGT	22	

RESULT 5  
US-09-925-065A-516005  
; Sequence 516005, Application US/09925065A





Qy 2055 AAGCCCGCGCGCGCCAGCAGAGAGCTGATTCGCCGCCCTGTGTGAGATGAGAGA 2114  
Db 63034 AGGACGACGAGAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGC 62975  
Qy 2115 TCGAGCGCTCGGCGCTGAGCCGGAATGAGAGAGATGTGTGAGAGAGCTGA 2166  
Db 62974 AGGAGGAGAGAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAG 62923

RESULT 10  
US-09-925-065A--516003/C  
; Sequence 516003, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 516003  
; LENGTH: 302  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-516003

Query Match 1.3%; Score 112.4; DB 6; Length 302;  
Best Local Similarity 99.1%; Pred. No. 6.3e-13;  
Matches 113; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 2214 GCCCAGCACTGTCAACAACAGCTCAGACCCGAGAGATCCCTCTCTCTCACTGAG 2273  
Db 114 GCACAGCACTGTCAACAACAGCTCAGACCCGAGAGATCCCTCTCTCTCACTGAG 55  
Qy 2274 CCGCCAGAGACACAGGCGAGATGGGCCCAAGCCCCCAGCCACCTTGGGCGCG 2327  
Db 54 CCGCCAGAGACACAGGCGAGATGGGCCCAAGCCCCCAGCCACCTTGGGCGCG 1

RESULT 11  
US-11-121-086-5  
; Sequence 5, Application US/1121086  
; Publication No. US20050266459A1  
; GENERAL INFORMATION:  
; APPLICANT: NIELSEN, KRISTEN V.  
; APPLICANT: POULSEN, TIM S.  
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES  
; FILE REFERENCE: 09138, 6000-00000  
; CURRENT APPLICATION NUMBER: US/11/121,086  
; PRIOR FILING DATE: 2005-05-04  
; PRIOR APPLICATION NUMBER: 60/567,570  
; PRIOR FILING DATE: 2004-05-04  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 5  
; LENGTH: 153376  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-121-086-5

Query Match 1.2%; Score 98.6; DB 12; Length 153376;

Best Local Similarity 44.9%; Pred. No. 7.8e-10;  
Matches 577; Conservative 0; Mismatches 699; Indels 9; Gaps 5;  
Qy 1417 GACTAAGAAAGATGAACTAATTAAGAGCTGTGTGAGACGAGCTTATCGGCGCCGCGCA 1476  
Db 15043 GAGGAGGAGAAACAGAGCTGTAGACACTGTGGGACCTCCAGATGCCGGGCACTG 15102  
Qy 1477 GAGCCAGCAGCAACAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1536  
Db 15103 GAGACAGAGAGAGAGAGAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15162  
Qy 1537 GCCCCGACAGCCAGGAGAGAAAGATGAGAAAGAGAAAGAGAGAGAGAGAGAGAG 1596  
Db 15163 GAGAGAACAG 15222  
Qy 1597 GAGAGAGAGCCGAGGTGTGAGAAACGACAGAGAAACCTCTCAAGAGAGAGAGAGAG 1656  
Db 15223 GAAACAG 15282  
Qy 1657 CACCTCAGGGAGAGCAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1715  
Db 15283 GAG 15342  
Qy 1716 CCAACAGCCAGGAAAGAGCAAGGCCCATCAACCCGCTCAATGCTTAATAGGCCAACA 1775  
Db 15343 GAAACAG 15402  
Qy 1776 GCGAGAGAGCCATCAAC-CCCCAGCAGAGCCGAGCTGTGCTTCAATGAGCTGAATGAG 1834  
Db 15403 GGAATCCCAAGTGTGAG 15462  
Qy 1835 AGTTCTCGCTGAGACAGAAAGAAATGAGAAACAGCCAGAAAGGTCTTCTGAAACAGCG 1894  
Db 15463 GAG 15522  
Qy 1895 GCACATGCTGTGCGCATGTG-CCCCGATGTGGGCTCCAAAGTGTGTGCACTGTAAAGA 1953  
Db 15523 CCAGATGCGGGGCGCTGTGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 15582  
Qy 1954 CTTTACTTCAACTTCAAGAAAGAGCAGAGCCTGATGAGATCTTGACAGCAGCAAGCT 2013  
Db 15583 ACAG 15642  
Qy 2014 GAAGATGAG 2073  
Db 15643 GAG 15702  
Qy 2074 CGAGAGGCTGCAATTCGCCGCCGTGTGTGAGATGAGATGAGAGCTGTGGGCTGAG 2133  
Db 15703 GAG 15762  
Qy 2134 CGGAAATGAG 2193  
Db 15763 ACAG 15822  
Qy 2194 GCCCAGAGAGAGATGAT 2253  
Db 15823 GGTGTAGAGACACTGTGGGAGCTGTCCAGATGTCCAGGGGCACTGTGAGACAGAGAGAGAG- 15881  
Qy 2254 CCGCTTCTCTCACTGAGGCGGCGCAGAGACAGAGGAGAGAGAGAGAGAGAGAGAGAG 2313  
Db 15882 GAG 15941  
Qy 2314 CACCTGGGCGCCGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2373  
Db 15942 AG 16001  
Qy 2374 GAGCCCATTTAGAGCCCAACCGGCGCTTGAAGCAGCCGAGGCGCTTACGCCCCACAGC 2433  
Db 16002 AG 16061  
Qy 2434 ACCCCATGCGCTCTGTGACCTCTCTGTGTGTCCCAAGAGAGAGAGAGAGAGAGAGAG 2493



```
/ CURRENT APPLICATION NUMBER: US/11/072,512
/ CURRENT FILING DATE: 2005-03-07
/ PRIOR APPLICATION NUMBER: US 60/350,978
/ PRIOR FILING DATE: 2002-01-25
/ PRIOR APPLICATION NUMBER: JP 2001-379298
/ NUMBER OF SEQ ID NOS: 4096
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1666
/ LENGTH: 2093
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-11-072-512-1666
```

```
Query Match      1.0%; Score 82.4; DB 9; Length 2093;
Best Local Similarity 47.1%; Pred. No. 6.9e-07;
Matches 356; Conservative 0; Mismatches 391; Indels 9; Gaps 3;

QY 936 AGGTGAGCGCATTCGAAACAAACCCGCGCGGCGCAAGAGCAAGTGCAGT 995
    |||||
DB 997 AGGAGGAAAGATGTGAGACAGAGAAAGCTCGAGAGCAGAGAGAGAGTGCAG 1056
    |||||
QY 996 ACTACGAAAGAGTTCCTGATCCGAGCAGCGGAGCTGAGAGCGCATGCGA 1055
    |||||
DB 1057 AGCAGGAGAGAGCTCGCGGAGCAGAAAGCTCGGAGCAGAGAGCAGATGCA 1116
    |||||
QY 1056 GCAGGCTGCGCAGCGGCGCAGTGGCTGTCCATG--TCGCGCGCGCCAGCGACAG 1113
    |||||
DB 1117 AGCAGAGAGAGAGATGTGGAGCAGAGAGAGAGATCGGGAGAGAGAGAGATGT 1176
    |||||
QY 1114 GGTGTCAAGATCATGATGCGCTTCA-GAGCAGAGAACTTGAAGAGCAGATGCC 1172
    |||||
DB 1177 GAGAGCAGAGAGAGCGCTGTGGGAGCAGAGAGACAGATGCGGAGCAGAGCAAGA 1236
    |||||
QY 1173 AGCTGCGCGTATCCCGCCATGCTGTACAGCTGACCAAGAGCGCATCAATTCA 1232
    |||||
DB 1237 TCGCGGACCAAGAGAGAGATGTGGAGCAGAGAGCTTACGGAGAGAGAGAG 1296
    |||||
QY 1233 ACATGAACGGGCTTATGCGCCGACCCCATGAGGTGTAACAAGACCGCAGTCATGA 1292
    |||||
DB 1297 GGATGCGGAGAGAGAGAGATGTGGAGCAGAGTGGAGAGATGCGGAGAGAGAG 1356
    |||||
QY 1293 TGTGAGTGAAGAGAGAGAGACCTTTCGGGAGAGATTCATGACATCCAGAACT 1352
    |||||
DB 1357 TGCAGAGACAGAGAGAGAGAGCGCGGACAGAGAGAGATGCAAGAGAGAGAG 1416
    |||||
QY 1353 TTGCGCTGATGCATCATTTCTGAGAGAGAGACAGTGGCTGAGTGCCTTATTA 1412
    |||||
DB 1417 TACGGGAGCGGAGAA-----GAGATGCGGAGAGAGAGAGATGCGGAGCAG 1470
    |||||
QY 1413 ACTGACTTAAGAAAGATGAACTTAAGAGCTGTGAGACGAGCTATCGCGCGCG 1472
    |||||
DB 1471 AGGAGAGATGAGAGAGAGAGAGAGATATGTGGAGCAGAGAGAGAGAGATGCG 1530
    |||||
QY 1473 GCAAGAGCCAGCAGCAACAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1532
    |||||
DB 1531 AGCAGAGGCTGCGGAGACAGAGAGAGCTGTGGAACAGAGAGAGATGCGAGCAG 1590
    |||||
QY 1533 CCATGCCCCGAGCAGCAGCAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAG 1592
    |||||
DB 1591 AGGAGAGATTTGGAGCAGAGAGAGAGAGATACGAGACAGAGAGATGTGGGAG 1650
    |||||
QY 1593 AGGAGAGAGAGAGCCGAGGTGTGAGAAAGACAGAGAGAGAGCTCTCAAGAGAG 1652
    |||||
DB 1651 AGAGAGAGATGTGCGAGAGAGAGAGATGTGGGCGCAGAGAGAGATGCGAGAG 1710
    |||||
QY 1653 ACGACACTCAGGGGAGAGCAACGAGAGAGAGAG 1688
    |||||
DB 1711 AGGAGAGATACAGATCAGAAAGAGAGATGCA 1746
    |||||
```

RESULT 15  
US-11-136-527-7472/c

```
/ Sequence 7472, Application US/1136527
/ Publication No. US20050287570A1
/ GENERAL INFORMATION:
/ APPLICANT: Wyeth
/ APPLICANT: Mounts, William M
/ TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
/ FILE REFERENCE: 031896-041000 (AM101086)
/ CURRENT APPLICATION NUMBER: US/11/136,527
/ CURRENT FILING DATE: 2005-05-25
/ PRIOR APPLICATION NUMBER: US 60/574,294
/ PRIOR FILING DATE: 2005-05-26
/ NUMBER OF SEQ ID NOS: 362830
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 7472
/ LENGTH: 1400
/ TYPE: DNA
/ ORGANISM: Rattus norvegicus
US-11-136-527-7472
```

```
Query Match      0.9%; Score 81.2; DB 12; Length 1400;
Best Local Similarity 59.6%; Pred. No. 1.1e-06;
Matches 137; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 1413 ACTGACTTAAGAGATGAGAACTTAAGAGCTGTGAGACGAGCTATCGCGCCCG 1472
    |||||
DB 507 AGCAGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 448
    |||||
QY 1473 GCAAGAGCCAGCAGCAACAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1532
    |||||
DB 447 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 388
    |||||
QY 1533 CCATGCCCCGAGCAGCAGAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAG 1592
    |||||
DB 387 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 328
    |||||
QY 1593 AGGAGAGAGAGAGCCGAGGTGTGAGAAAGACAGAGAGAGAGCTCTCAAG 1642
    |||||
DB 327 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 278
    |||||
```

Search completed: March 12, 2006, 15:36:41  
Job time : 2101 secs



GenCore version 5.1.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2006, 10:17:44 ; Search time 5962 Seconds  
(without alignments)  
11888.098 Million cell updates/sec

Title: US-09-522-753-4

Perfect score: 8564.8  
Sequence: 1 catgctggcgcacacagc.....caaaaaaaaaaaaaa 8571

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 413468905 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8561	99.7	8561	6 US-10-174-014-11	Sequence 11, App1
2	8541	99.7	8561	10 US-11-052-344-1	Sequence 1, App1
3	8459.2	98.8	8667	5 US-10-087-192-653	Sequence 653, App
4	8324.6	97.2	8686	3 US-09-819-104A-1	Sequence 1, App1
5	8324.6	97.2	8686	6 US-10-174-014-4	Sequence 4, App1
6	8324.6	97.2	8686	8 US-10-723-860-1739	Sequence 1739, App
7	8222.2	96.0	9079	8 US-10-723-860-6114	Sequence 6114, App
8	7554	88.2	7521	6 US-10-414-692-14	Sequence 14, App1
9	7334.6	85.6	7521	3 US-09-819-104A-3	Sequence 3, App1
10	5037.2	58.8	8544	9 US-10-087-192-652	Sequence 4, App1
11	5037.2	58.8	8544	9 US-10-764-420-1526	Sequence 1526, App
12	4797	56.0	7386	3 US-09-819-104A-6	Sequence 6, App1
13	4543	53.0	7534	5 US-10-087-192-650	Sequence 650, App
14	2618.6	30.6	2930	5 US-10-146-473-14	Sequence 14, App1
15	2618.6	30.6	2930	6 US-10-174-014-13	Sequence 13, App1
16	1138.6	13.3	23380	6 US-10-087-192-652	Sequence 652, App
17	1132.4	13.2	221000	6 US-10-174-014-12	Sequence 12, App1
18	949.4	11.1	1027	3 US-09-925-297-9	Sequence 9, App1
19	898.4	10.5	956	3 US-09-887-527-42	Sequence 42, App1
20	898.4	10.5	956	7 US-10-796-174-42	Sequence 42, App1
21	630	7.4	7940	6 US-10-341-434-235	Sequence 235, App
22	630	7.4	7940	8 US-10-473-974-191	Sequence 191, App
23	630	7.4	7940	9 US-10-887-553A-664	Sequence 664, App

24	630	7.4	7940	9 US-10-956-157-1580	Sequence 1580, App
25	630	7.4	7940	9 US-10-745-237-221	Sequence 221, App
26	626.8	7.3	2745	6 US-10-414-692-15	Sequence 15, App1
27	553.6	6.5	692	5 US-10-027-632-102554	Sequence 102554, App
28	553.6	6.5	692	5 US-10-027-632-102555	Sequence 102555, App
29	553.6	6.5	692	6 US-10-027-632-102554	Sequence 102554, App
30	553.6	6.5	692	6 US-10-027-632-102555	Sequence 102555, App
31	542.4	6.3	718	8 US-10-278-698-515	Sequence 515, App
32	542.4	6.3	718	8 US-10-278-698-1031	Sequence 1031, App
33	432.8	5.1	91141	5 US-10-087-192-649	Sequence 649, App
34	409.4	4.8	427	9 US-10-779-543-12895	Sequence 12895, App
35	376	4.4	872	9 US-10-779-543-8757	Sequence 8757, App
36	370	4.3	381	3 US-09-918-995-7653	Sequence 7653, App
37	314.4	3.7	458	3 US-09-918-995-20064	Sequence 20064, App
38	268.6	3.1	520	3 US-09-920-300A-485	Sequence 485, App
39	268.6	3.1	520	5 US-10-033-528-485	Sequence 485, App
40	268.6	3.1	520	6 US-10-099-926-485	Sequence 485, App
41	268.6	3.1	520	6 US-10-961-527-485	Sequence 485, App
42	258.4	3.0	3969	6 US-10-006-285-471	Sequence 471, App
43	211.4	2.5	499	3 US-09-918-995-19696	Sequence 19696, App
44	191.4	2.2	390	6 US-10-355-716-86	Sequence 86, App1
45	187.8	2.2	789	9 US-10-779-543-5226	Sequence 5226, App

## ALIGNMENTS

RESULT 1  
US-10-174-014-11  
; Sequence 11, Application US/10174014  
; Publication No. US20040005292A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Susan M. Freiler  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION  
; FILE REFERENCE: PTS-0012  
; CURRENT APPLICATION NUMBER: US/10/174,014  
; CURRENT FILING DATE: 2002-06-17  
; NUMBER OF SEQ ID NOS: 73  
; SEQ ID NO 11  
; LENGTH: 8561  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)...(7555)  
US-10-174-014-11

Query Match Best Local Similarity 99.7%; Score 8541; DB 6; Length 8561;  
Matches 8560; Conservative 1; Mismatches 0; Indels 10; Gaps 1;

QY	1	CATGTCGGGCTCCACACAGCTTGAGCAGAGAGGCGCACTGAGCCCGCTAACCC	60
DB	1	CATGTCGGGCTCCACACAGCTTGAGCAGAGAGGCGCACTGAGCCCGCTAACCC	60
QY	61	GGCCCAAGCCCTTCTTACCCAGTGCAGATCGCCGAGACAGCAGAGTGGGGCTCCT	120
DB	61	GGCCCAAGCCCTTCTTACCCAGTGCAGATCGCCGAGACAGCAGAGTGGGGCTCCT	120
QY	121	GGAGTACAGACCACTCCGCGAGATGCTCCCACTGCGCGGGCTCATCATCCA	180
DB	121	GGAGTACAGACCACTCCGCGAGATGCTCCCACTGCGCGGGCTCATCATCCA	180
QY	181	GGCCCAAGCCCGAGAGCCCTCTGCTGTAGTTCAGCCCGGAGTGAAGGTTCCCA	240
DB	181	GGCCCAAGCCCGAGAGCCCTCTGCTGTAGTTCAGCCCGGAGTGAAGGTTCCCA	240
QY	241	GGAGTTCACCTGCGGCGCAGAGTTCCTCATATCTGCGGAGTGGGAATGACAGAT	300
DB	241	GGAGTTCACCTGCGGCGCAGAGTTCCTCATATCTGCGGAGTGGGAATGACAGAT	300

QY	301	GGAGTTCAATTAAAGCAAGCGCCCTCGGCTTGAAGTGCTGCTGA	CCCCCTGCTGGCAC	360
Db	301	GGAGTTCAATTAAAGCAAGCGCCCTCGGCTTGAAGTGCTGCTGA	CCCCCTGCTGGCAC	360
QY	361	GTCACCCCTGCTGGGCAAGCGGCAAGCTGCGGAGCTTGAAGCACTCA	CCAGAGACCTGAG	420
Db	361	GTCACCCCTGCTGGGCAAGCGGCAAGCTGCGGAGCTTGAAGCACTCA	CCAGAGACCTGAG	420
QY	421	CTTGAAGGCAAGCTGGAACCGGTGTCTCCCCCAGCCCCCGCACTGA	CCCTGAGCT	480
Db	421	CTTGAAGGCAAGCTGGAACCGGTGTCTCCCCCAGCCCCCGCACTGA	CCCTGAGCT	480
QY	481	GGAGTGTGTGCGGCCACAGGCTGTCCAAAGAGAGCTGATCCAGAA	CAATGACCGGTGTGA	540
Db	481	GGAGTGTGTGCGGCCACAGGCTGTCCAAAGAGAGCTGATCCAGAA	CAATGACCGGTGTGA	540
QY	541	CCGAGAGATTCACATGTGTAGAGACAGATCTCTAAGCTGAAGAA	GAGACGCAACAGCT	600
Db	541	CCGAGAGATTCACATGTGTAGAGACAGATCTCTAAGCTGAAGAA	GAGACGCAACAGCT	600
QY	601	GGAAGAGAGGCTGCAAGCCGCCAGACTTGAAGAGCCGCTGTCA	CCGCCCATGGA	660
Db	601	GGAAGAGAGGCTGCAAGCCGCCAGACTTGAAGAGCCGCTGTCA	CCGCCCATGGA	660
QY	661	GTCGAAGACCCGCAAGCTGTGTGCAGATCTACACAGAAACCGGA	AGAAAGGCTGAAGC	720
Db	661	GTCGAAGACCCGCAAGCTGTGTGCAGATCTACACAGAAACCGGA	AGAAAGGCTGAAGC	720
QY	721	TGCACATCGAATTCGAAAGGCTTGAGGCCCCAGGTGAGAGCTGC	CGCTGTCAACAGACC	780
Db	721	TGCACATCGAATTCGAAAGGCTTGAGGCCCCAGGTGAGAGCTGC	CGCTGTGTCAACAGACC	780
QY	781	CTCCGACACCCGGCAGTATCATAGAAATCAAAATAAACAGCGGAT	TCGGAAGAAAGCT	840
Db	781	CTCCGACACCCGGCAGTATCATAGAAATCAAAATAAACAGCGGAT	TCGGAAGAAAGCT	840
QY	841	AATCTTGTACTTCAAGAGAGAGAAATCAAGCTTCGAAACAAAT	TGAAAGCAAGATTCTGCA	900
Db	841	AATCTTGTACTTCAAGAGAGAGAAATCAAGCTTCGAAACAAAT	TGAAAGCAAGATTCTGCA	900
QY	901	GCGCTATACACAGCTCATGAGAGCCCTTGAAAAAAAAGGTGAG	CGGCATTCGAAACACC	960
Db	901	GCGCTATACACAGCTCATGAGAGCCCTTGAAAAAAAAGGTGAG	CGGCATTCGAAACACC	960
QY	961	GCGCCGCGGGCCAAAGAGCAAGGTGCGGAGTCTACAAAGCAG	ATTCCCTGAGAT	1020
Db	961	GCGCCGCGGGCCAAAGAGCAAGGTGCGGAGTCTACAAAGCAG	ATTCCCTGAGAT	1020
QY	1021	CCGCAAGCAGCGCGAGCTTCAGAGCGCATGACAGAGGCTGGCC	CACCGGCGCATGTGG	1080
Db	1021	CCGCAAGCAGCGCGAGCTTCAGAGCGCATGACAGAGGCTGGCC	CACCGGCGCATGTGG	1080
QY	1081	GCTGTCCATGTCCGCGCGCCGAGGACACAGAGGTGCAGAGAT	CAATGATGAGCTCTTC	1140
Db	1081	GCTGTCCATGTCCGCGCGCCGAGGACACAGAGGTGCAGAGAT	CAATGATGAGCTCTTC	1140
QY	1141	AGAGCAGAGAACTTGAGAAAGAGATCGCCAGCTGATCCCGCCAT	CTGCTGTA	1200
Db	1141	AGAGCAGAGAACTTGAGAAAGAGATCGCCAGCTGATCCCGCCAT	CTGCTGTA	1200
QY	1201	CGAGCGTGAACAAGCGCATCAAGTTCAATCAATGAAACGAGGCT	TTATGCGCACCCAT	1260
Db	1201	CGAGCGTGAACAAGCGCATCAAGTTCAATCAATGAAACGAGGCT	TTATGCGCACCCAT	1260
QY	1261	GAAAGTGTAAACAAGCCGCGCAGGTCAATGAACATGTGAGTA	GACGAGAAAGACATTT	1320
Db	1261	GAAAGTGTAAACAAGCCGCGCAGGTCAATGAACATGTGAGTA	GACGAGAAAGACATTT	1320
QY	1321	CCGGGAGAAATTCATGACAGCATCCCAAGAACTTTGGCTGTAT	CGCATATTTCTGAGAG	1380
Db	1321	CCGGGAGAAATTCATGACAGCATCCCAAGAACTTTGGCTGTAT	CGCATATTTCTGAGAG	1380
QY	1381	GAAGCAGTGGCTGAGTGCCTCTCTATTACTGACTGA	CTAAGAAAGATGAACTATTA	1440

[illegible]

D	b	2461	1GTGTCCCCCAAGAGAGAGAAAGAGAGAGAAACCGACACGCCCCCACTGAGAGAGG	2520
O	y	2521	GGAGGAGCAGAAAGCCCCCGCGCTGAGAGACTGGCACTGGACAAGGAAAGCCGAGA	2580
D	b	2521	GGAGGAGCAGAAAGCCCCCGCGCTGAGAGACTGGACAAGGAAAGCCGAGA	2580
O	y	2551	GGCCCTCAAGAGCCAGGTGCACGAGAGAAAGCCGAGAGAGGGGCGCGCAAGGCAAGACG	2640
D	b	2581	GGCCCTCAAGAGCCAGGTGCACGAGAGAAAGCCGAGAGAGGGGCGCGCAAGGCAAGACG	2640
O	y	2641	GGAGGCGCTGAGGCGCAGCGCCGAGGGGGCGCTCAAGGCAGAGAAAGAGAGGGCGGAG	2700
D	b	2641	GGAGGCGCTGAGGCGCAGCGCCGAGGGGGCGCTCAAGGCAGAGAAAGAGAGGGCGGAG	2700
O	y	2701	CGGCAAGGCGCAACAATGCGAAGGCTCGGGCGCCCCCGGAGCAGACGACTTCAGTGTAC	2760
D	b	2701	CGGCAAGGCGCAACAATGCGAAGGCTCGGGGGCCCCCGGAGCAGACGACTTCAGTGTAC	2760
O	y	2761	CTGCAGTGCACACGAGGTGATGAGGCGGAGGGCGGCGACAAGAACCGGCTGTGTCCC	2820
D	b	2761	CTGCAGTGCACACGAGGTGATGAGGCGGAGGGCGGCGACAAGAACCGGCTGTGTCCC	2820
O	y	2821	AAAGGCGCAGCCCTCCACCCCGGACCTGGGCGACCCCGGGCCAAATGCTCAACCCGAAAGCC	2880
D	b	2821	AAAGGCGCAGCCCTCCACCCCGGACCTGGGCGACCCCGGGCCAAATGCTCAACCCGAAAGCC	2880
O	y	2881	ACTGGACCTGAAGCAGCTGAAGACGACGAGCGGCTGCATCCCCCCCAATCCAGTCACAA	2940
D	b	2881	ACTGGACCTGAAGCAGCTGAAGACGACGAGCGGCTGCATCCCCCCCAATCCAGTCACAA	2940
O	y	2941	AGTCCATAGCCCCCCCCGGGAGAGACGACGCTCCACCAAGCAGCTCCCGACGCCAAC	3000
D	b	2941	AGTCCATAGCCCCCCCCGGGAGAGACGACGCTCCACCAAGCAGCTCCCGACGCCAAC	3000
O	y	3001	GCCACCGGAAAACCTGCGACCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3060
D	b	3001	GCCACCGGAAAACCTGCGACCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	3060
O	y	3061	GGGCAAGAGCAGAGAGCCGCGACACCCCGCGCGACAGAGAGAGCTTCGACGCGAGGCCA	3120
D	b	3061	GGGCAAGAGCAGAGAGCCGCGACACCCCGCGCGACAGAGAGAGCTTCGACGCGAGGCCA	3120
O	y	3121	GAAAGCTGCTGGGGAGCCCCCTTGCTGGACTTCGGGCTGCGCTTCCTCCCGTGCCTCCG	3180
D	b	3121	GAAAGCTGCTGGGGAGCCCCCTTGCTGGACTTCGGGCTGCGCTTCCTCCCGTGCCTCCG	3180
O	y	3181	TGAGGTGATCAAGGCTCTCCCGCATATGCCCGGAAACCTCAGCTTCTCTACGCTCCAC	3240
D	b	3181	TGAGGTGATCAAGGCTCTCCCGCATATGCCCGGAAACCTCAGCTTCTCTACGCTCCAC	3240
O	y	3241	TGGTCAACCACTGGCCCTGGGCGCTCCATGACATGACCTGGCCGCGCTCTGGCCGCGCAC	3300
D	b	3241	TGGTCAACCACTGGCCCTGGGCGCTCCATGACATGACCTGGCCGCGCTCTGGCCGCGCAC	3300
O	y	3301	CACCATCTCAACCCGCGCTCCCTCATCTCTCTGCAAGACACCCAGCGTCTCGAGAG	3360
D	b	3301	CACCATCTCAACCCGCGCTCCCTCATCTCTCTGCAAGACACCCAGCGTCTCGAGAG	3360
O	y	3361	GCAAAATAGTGCATCTCTCCAAAGAAATGTGCGTCAAGCTCCACGCTCCCGTACTCAGAGCA	3420
D	b	3361	GCAAAATAGTGCATCTCTCCAAAGAAATGTGCGTCAAGCTCCACGCTCCCGTACTCAGAGCA	3420
O	y	3421	TGGCAAGGCGCCGGTGGGCGCTGTCAACATGGGGGTGCGCTCGCCATATGAGACCCCAAAA	3480
D	b	3421	TGGCAAGGCGCCGGTGGGCGCTGTCAACATGGGGGTGCGCTCGCCATATGAGACCCCAAAA	3480
O	y	3481	GCTGGCACCTTCAAGCGGAGTGAAGCAGAGCAGACTGTCCCAAGGGGCGACAGGCTGGGCC	3540
D	b	3481	GCTGGCACCTTCAAGCGGAGTGAAGCAGAGCAGACTGTCCCAAGGGGCGACAGGCTGGGCC	3540
O	y	3541	ACCGGAGAGCTTGGGGGTGGCCACAAGCCAGAGAGGCGTCCGTGTGAGAGGAGACAGTCT	3600
D	b	3541	ACCGGAGAGCTTGGGGGTGGCCACAAGCCAGAGAGGCGTCCGTGTGAGAGGAGACAGTCT	3600

QY	3601	GGGCTCAGTTC	CCGGGCGGAAAGCAT	CACCAAGGCA	TTCCAGCA	CACGGGTG	CCCTCGGA	3660
Db	3601	GGGCTCAGTTC	CCGGGCGGAAAGCAT	CAACAAAGGCA	TTCCAGCA	CACGGGTG	CCCTCGGA	3660
QY	3661	CAGGCGCAT	CATACATCCGGGCT	CAATCA	CCCAAGGC	ACGGCA	CTGACGTCTGTACAA	3720
Db	3661	CAGGCGCAT	CAATACCGGGCT	CAATCA	CCCAAGGC	ACGGCA	CTGACGTCTGTACAA	3720
QY	3721	GGGCAACAT	CACCAAGATCAT	TCGGCGAGAG	CAGCGCGGAT	GGCTTTGAC	CCCGCGCCGGGA	3780
Db	3721	GGGCAACAT	CACCAAGATCAT	TCGGCGAGAG	CAGCGCGGAT	GGCTTTGAC	CCCGCGCCGGGA	3780
QY	3781	GGA	CAGCCTTG	CCCAAGGGGCA	CGTCA	TCA	GAAAGGCA	3840
Db	3781	GGA	CAGCCTTG	CCCAAGGGGCA	CGTCA	TCA	GAAAGGCA	3840
QY	3841	TGA	GGGTGG	CAATGTCT	TGACCC	CAGTGTCT	CCAAAGAGACGGCA	3900
Db	3841	TGA	GGGTGG	CAATGTCT	TGACCC	CAGTGTCT	CCAAAGAGACGGCA	3900
QY	3901	CCCCCAT	GAGA	CGGCGCGCCCCCA	AGCGCA	CCCTA	GACATGA	3960
Db	3901	CCCCCAT	GAGA	CGGCGCGCCCCCA	AGCGCA	CCCTA	GACATGA	3960
QY	3961	AGC	CACTCTCT	CAGCCAG	CATCGA	AGGTCT	CATG	4020
Db	3961	AGC	CACTCTCT	CAGCCAG	CATCGA	AGGTCT	CATG	4020
QY	4021	CAG	CCCCCA	CACTCA	AAAGAC	CA	CATCCGCGGT	4080
Db	4021	CAG	CCCCCA	CACTCA	AAAGAC	CA	CATCCGCGGT	4080
QY	4081	TCGGT	CTTA	GTGAGG	CA	CAGAGG	CA	4140
Db	4081	TCGGT	CTTA	GTGAGG	CA	CAGAGG	CA	4140
QY	4141	GGA	GGGCA	CA	CGCTCC	CGGCCCCCA	CGGCTCA	4200
Db	4141	GGA	GGGCA	CA	CGCTCC	CGGCCCCCA	CGGCTCA	4200
QY	4201	GGC	CTTG	GGCCCT	CTGA	AGCTGA	AGCGG	4260
Db	4201	GGC	CTTG	GGCCCT	CTGA	AGCTGA	AGCGG	4260
QY	4261	GG	CGGCG	CGCTCA	TCCATGA	AGTCCCG	GAGAG	4320
Db	4261	GG	CGGCG	CGCTCA	TCCATGA	AGTCCCG	GAGAG	4320
QY	4321	CTT	GGCCCCCG	CGCTCA	AGAGG	AGCTCAT	CA	4380
Db	4321	CTT	GGCCCCCG	CGCTCA	AGAGG	AGCTCAT	CA	4380
QY	4381	CAC	CGGCG	CGTCA	CACTG	GGCTT	CA	4440
Db	4381	CAC	CGGCG	CGTCA	CACTG	GGCTT	CA	4440
QY	4441	CGG	CGGCA	CGTTC	CA	AGGCTT	CA	4500
Db	4441	CGG	CGGCA	CGTTC	CA	AGGCTT	CA	4500
QY	4501	ACG	GTGCTG	CTA	CGAGAG	AGCTCT	GA	4560
Db	4501	ACG	GTGCTG	CTA	CGAGAG	AGCTCT	GA	4560
QY	4561	CTC	CA	TTGCGCG	CGCGCC	CGCTCA	TTG	4620
Db	4561	CTC	CA	TTGCGCG	CGCGCC	CGCTCA	TTG	4620
QY	4621	CTG	ACCTTA	GA	AGACCA	CGGGG	CA	4680
Db	4621	CTG	ACCTTA	GA	AGACCA	CGGGG	CA	4680

QY 4681 GAACATGCGGAGCCCAACGCGCGCTGCGAGAGGGAGCCCTTTGCTGCACGAAGCATC 4740  
Db 4681 GACCATGCGGGAGCCCAACGCGCGCTGCGAGAGGGAGCCCTTTGCTGCACGAAGCATC 4740  
QY 4741 CCAAGACCGGAAAGCTGAGCGCTCTGTGAGATTCGCGAAGTCCCGGACAGACCGGT 4800  
Db 4741 CCAAGACCGGAAAGCTGAGCGCTCTGTGAGATTCGCGAAGTCCCGGACAGACCGGT 4800  
QY 4801 GCCCGAGACACCAACCAACCCCATCTCGCCCTATAGACACTGCTTGGGGGCTGAGTGG 4860  
Db 4801 GCCCGAGACACCAACCAACCCCATCTCGCCCTATAGACACTGCTTGGGGGCTGAGTGG 4860  
QY 4861 CGTGAACCTGTATCGCAGCCACATCCCTGCGCTTGAACCCCACTCCATACCCCGGG 4920  
Db 4861 CGTGAACCTGTATCGCAGCCACATCCCTGCGCTTGAACCCCACTCCATACCCCGGG 4920  
QY 4921 CATCCCTCTGAGACGACGCGCTGCTTACTACTGCGCCGACACTGGGCCCCCAACCCAC 4980  
Db 4921 CATCCCTCTGAGACGACGCGCTGCTTACTACTGCGCCGACACTGGGCCCCCAACCCAC 4980  
QY 4981 CTACCCGACCTGTATCCCAACCCCTACCTCATTCGCGGGCTTACCCCGACGCGCGCTGGA 5040  
Db 4981 CTACCCGACCTGTATCCCAACCCCTACCTCATTCGCGGGCTTACCCCGACGCGCGCTGGA 5040  
QY 5041 GAAACCGGACAGCATCATCAATGACTATACATCACTTCGACAGAGATGACCAACAACGCG 5100  
Db 5041 GAAACCGGACAGCATCATCAATGACTATACATCACTTCGACAGAGATGACCAACAACGCG 5100  
QY 5101 CACCGCCATTTGGGCCCGACAGCTGATATGTGAGGGGGCTCTTGCGCCCGGAGTCTTGCT 5160  
Db 5101 CACCGCCATTTGGGCCCGACAGCTGATATGTGAGGGGGCTCTTGCGCCCGGAGTCTTGCT 5160  
QY 5161 GGCACTCAACTGCGTGCGGGTCCCGAGAGGCATCATGCACTGTCCAAAGTSCACACT 5220  
Db 5161 GGCACTCAACTGCGTGCGGGTCCCGAGAGGCATCATGCACTGTCCAAAGTSCACACT 5220  
QY 5221 GCTGTGCTGCTGTCGCCGACACAGGACACCCAGCCACCGGCATGAAACCGCTTGCTTA 5280  
Db 5221 GCTGTGCTGCTGTCGCCGACACAGGACACCCAGCCACCGGCATGAAACCGCTTGCTTA 5280  
QY 5281 CCTCCCAACCGGGGCCCGACCTTCAAGACGCGCCACAGCACTCTCCCACTTCCCGAG 5340  
Db 5281 CCTCCCAACCGGGGCCCGACCTTCAAGACGCGCCACAGCACTCTCCCACTTCCCGAG 5340  
QY 5341 AGGTCCAACAACCTTGAACAACCAACAACGCTCTGCTCGAGCGGGAGCGAGACCG 5400  
Db 5341 AGGTCCAACAACCTTGAACAACCAACAACGCTCTGCTCGAGCGGGAGCGAGACCG 5400  
QY 5401 GGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAATTCATCTCAAGTCCACACGAC 5460  
Db 5401 GGATCGAGAGCGGGACCGGGATCGGGAGCGGGAAAAATTCATCTCAAGTCCACACGAC 5460  
QY 5461 GGTGAGACACCGCATCTGTGAGACCTGTGTACAAGACAGACCGGAGAGAGGGGAG 5520  
Db 5461 GGTGAGACACCGCATCTGTGAGACCTGTGTACAAGACAGACCGGAGAGAGGGGAG 5520  
QY 5521 CAGCGCGGGGGGTGGGGGAGCAGACGCGCGCGCTCCACTCCCAATGACCAACAGA 5580  
Db 5521 CAGCGCGGGGGGTGGGGGAGCAGACGCGCGCGCTCCACTCCCAATGACCAACAGA 5580  
QY 5581 CTGCGCCATCTCTCCGAGCCCAAGATGCTCTTCAGACAGAGACCAAGTGTCTTCAAA 5640  
Db 5581 CTGCGCCATCTCTCCGAGCCCAAGATGCTCTTCAGACAGAGACCAAGTGTCTTCAAA 5640  
QY 5641 CACAGGCAATGAAGGGTATCATACCGCTGTGAGAGCCCAAGGCCCAAGTCTTGAAGT 5700  
Db 5641 CACAGGCAATGAAGGGTATCATACCGCTGTGAGAGCCCAAGGCCCAAGTCTTGAAGT 5700  
QY 5701 CACCTCAACTCTCTCAACCGCTTCCGCGAGCTGCAATTCCTGACCAACCACTGCGCC 5760  
Db 5701 CACCTCAACTCTCTCAACCGCTTCCGCGAGCTGCAATTCCTGACCAACCACTGCGCC 5760  
QY 5761 ACTGGGCGGCAACCTCGATGGGGTCTAACCTTACCTCATGAGACCGGTCTTGCCCAA 5820

Db 5761 ACTGGGCGGCAACCTCGATGGGGTCTAACCTTACCTCATGAGACCGGTCTTGCCCAA 5820  
QY 5821 GGAAGGCCCCCGGGGTGCGCCGGGACAGAGCGGGCCCGGACAGACCGGACATGCTTCT 5880  
Db 5821 GGAAGGCCCCCGGGGTGCGCCGGGACAGAGCGGGCCCGGACAGACCGGACATGCTTCT 5880  
QY 5881 CGCAGAGCCCCAGCGCGCTCCGGCTGAGACCGCGCTCTTCCCGGACAAAGGCTCGGA 5940  
Db 5881 CGCAGAGCCCCAGCGCGCTCCGGCTGAGACCGCGCTCTTCCCGGACAAAGGCTCGGA 5940  
QY 5941 GCCCGGCCCCCTAAGTGTCTCTGTCTCTGGGCAAGCCACATTCGCGGACCCCTGCGAA 6000  
Db 5941 GCCCGGCCCCCTAAGTGTCTCTGTCTCTGGGCAAGCCACATTCGCGGACCCCTGCGAA 6000  
QY 6001 GAACTGCACTTCAACAACGACGCGAGCCGAGCCGCGGACCACTGCTGAGCTCGGA 6060  
Db 6001 GAACTGCACTTCAACAACGACGCGAGCCGAGCCGCGGACCACTGCTGAGCTCGGA 6060  
QY 6061 CCGGACCGGGAAAAAGCTCAAGTAAACCTTTTCAATCCAGAGAACTGAACTCCGTTT 6120  
Db 6061 CCGGACCGGGAAAAAGCTCAAGTAAACCTTTTCAATCCAGAGAACTGAACTCCGTTT 6120  
QY 6121 TCTGGGTTTACAACGACAGAGCTTACAGCCCGGAGGGGTGAGCCCTGACCCCTGTAG 6180  
Db 6121 TCTGGGTTTACAACGACAGAGCTTACAGCCCGGAGGGGTGAGCCCTGACCCCTGTAG 6180  
QY 6181 CTCACCCGATGACCCACAGCAAGGGGCTCCCAAGACCTGGAAGACTGACAAAG 6240  
Db 6181 CTCACCCGATGACCCACAGCAAGGGGCTCCCAAGACCTGGAAGACTGACAAAG 6240  
QY 6241 CCACCTGAGGGGGAGCTGCGGCGCCCAAGACAGCGCCGCTGGAAGCTTGGCGGGAGGC 6300  
Db 6241 CCACCTGAGGGGGAGCTGCGGCGCCCAAGACAGCGCCGCTGGAAGCTTGGCGGGAGGC 6300  
QY 6301 CGCCCACTCCCAACCTGCGGCGCTGCTGAGAGCCAGCCCTGTGTCAAGCCGCTGCT 6360  
Db 6301 CGCCCACTCCCAACCTGCGGCGCTGCTGAGAGCCAGCCCTGTGTCAAGCCGCTGCT 6360  
QY 6361 CGAGACGCGCCCGAGGGGTCAAAAGGTACACAGGGGTGTCAACCCGAGCCAGACATGAG 6420  
Db 6361 CGAGACGCGCCCGAGGGGTCAAAAGGTACACAGGGGTGTCAACCCGAGCCAGACATGAG 6420  
QY 6421 TGAAGTATCAACAAGACTTACACCCCGGACCAACCAAGAGGCTCAAGCCCTGCGC 6480  
Db 6421 TGAAGTATCAACAAGACTTACACCCCGGACCAACCAAGAGGCTCAAGCCCTGCGC 6480  
QY 6481 CGCCCTCTTATCTTCTTCCCTGGGGCCAGCTGCCCTGTGAACTTCGCGCGCCACG 6540  
Db 6481 CGCCCTCTTATCTTCTTCCCTGGGGCCAGCTGCCCTGTGAACTTCGCGCGCCACG 6540  
QY 6541 CAGTGAACCTTACCTCCGCGCCCGGACCATGTGTCGCCGCGGTGACTCCCCCAAG 6600  
Db 6541 CAGTGAACCTTACCTCCGCGCCCGGACCATGTGTCGCCGCGGTGACTCCCCCAAG 6600  
QY 6601 CGAAGGGGCAAGAGGTCTCCAGAGCCCAACAAAGCTTCGCTTGGTGTGTGAGGA 6660  
Db 6601 CGAAGGGGCAAGAGGTCTCCAGAGCCCAACAAAGCTTCGCTTGGTGTGTGAGGA 6660  
QY 6661 CGGTATTGAACCTGTGTCTCCCAACGAGGGGCAATGACGAGGCACTCCCGAGTGC 6720  
Db 6661 CGGTATTGAACCTGTGTCTCCCAACGAGGGGCAATGACGAGGCACTCCCGAGTGC 6720  
QY 6721 TGTGTACCGCGTGTATCCGGGATGGGGAAACAGACGAGCCAGCAGAGATGGGCTCAA 6780  
Db 6721 TGTGTACCGCGTGTATCCGGGATGGGGAAACAGACGAGCCAGCAGAGATGGGCTCAA 6780  
QY 6781 GTCTCAGGCAACAACAGCAGCGCCAGCTTCTTCAAGAGCTGACCGAGGCAATC 6840  
Db 6781 GTCTCAGGCAACAACAGCAGCGCCAGCTTCTTCAAGAGCTGACCGAGGCAATC 6840  
QY 6841 CGCCATGCTCAAGTCAAGAAAGAGATCAACAAGAGCTGAACCCCAACCGGAA 6900

Dp	6841	CGCCATGCTCAAGTCAGAAAGCAAGAGATCAACAAGAAAGCTGAACAACCCACAACCGGAA	6900
Qy	6901	TGAGCTGTAAATACAAATATACAGCCAGCTTGGACGGAGATCTTTCAATATGCCCCCATAC	6960
Dp	6901	TGAGCTGTAAATACAAATATACAGCCAGCTTGGACGGAGATCTTTCAATATGCCCCCATAC	6960
Qy	6961	CGGAACAGGCTTTATGACCTTATAGAAAGCAGGCGGTGAGGAAACATGCAACACCAAT	7020
Dp	6961	CGGAACAGGCTTTATGACCTTATAGAAAGCAGGCGGTGAGGAAACATGCAACACCAAT	7020
Qy	7021	GGGGCTGAGAGCCATAATATTAGAAAGCAGCTATGGGTAAATATGACAGTGGGAAAGATC	7080
Dp	7021	GGGGCTGAGAGCCATAATATTAGAAAGCAGCTATGGGTAAATATGACAGTGGGAAAGATC	7080
Qy	7081	CCCGCCGCTCAGCGCCAAATGCTTTTAAACCTCTGAATGCAGTGCAGCCTGCCGCTGC	7140
Dp	7081	CCCGCCGCTCAGCGCCAAATGCTTTTAAACCTCTGAATGCAGTGCAGCCTGCCGCTGC	7140
Qy	7141	TATGCGCCATTAACCGTGTGTGACGGAACGGAATGACACACTCAGCTTCGCAAGGTGGCGG	7200
Dp	7141	TATGCGCCATTAACCGTGTGTGACGGAACGGAATGACACACTCAGCTTCGCAAGGTGGCGG	7200
Qy	7201	CGGGAAGGCAAGGCTCTTGGCAGACCCAGGAGCGGAAAGCCAAAGTCCCGGCCCCGGG	7260
Dp	7201	CGGGAAGGCAAGGCTCTTGGCAGACCCAGGAGCGGAAAGCCAAAGTCCCGGCCCCGGG	7260
Qy	7261	CCTGGCATCTTGGGAGACCGGACCACTCTGTCTCTCAGTGTGACTTCGAGGGAGACTGCA	7320
Dp	7261	CCTGGCATCTTGGGAGACCGGACCACTCTGTCTCTCAGTGTGACTTCGAGGGAGACTGCA	7320
Qy	7321	CCGCCGGAACCGCGCTCACCACCAACCGGCTGTGGAGAGACAGGCCCTGCTCGCAGGTTCCAC	7380
Dp	7321	CCGCCGGAACCGCGCGTCAACCAACCGGCTGTGGAGAGACAGGCCCTGCTCGCAGGTTCCAC	7380
Qy	7381	GGCAATCCCTCCATCAACCCCGCTGATCATGACGGGCTGAGGCGGGGTGCATAGGTTCCCAAC	7440
Dp	7381	GGCAATTCCTCCATCAACCCCGCTGATCATGACGGGCTGAGGCGGGGTGCATAGGTTCCCAAC	7440
Qy	7441	CCCAACCGGCGCTCCCGCGGAGCAGCGGAGCCCTCGCTGCGTGGCCCCCAACACGCTGGAGCA	7500
Dp	7441	CCCAACCGGCGCTCCCGCGGAGCAGCGGAGCCCTCGCTGCGTGGCCCCCAACACGCTGGAGCA	7500
Qy	7501	GGAGGCCCAAGCACTGCTCTGCTCGAGTACGAGACACTCTCCGACAGCGAGTGACTCAG	7560
Dp	7501	GGAGGCCCAAGCACTGCTCTGCTCGAGTACGAGACACTCTCCGACAGCGAGTGACTCAG	7560
Qy	7561	AAACAAGGCGGGGGGGGGGGCGGGCGGTGEBURBSBRUJTCAGGTTCCCAAGAGCCACAAGAA	7620
Dp	7561	AAACAAGGCGGGGGGGGGGGGGGGCGG-----TGTCAAGGTCCAGAGAACCAAGAA	7610
Qy	7621	CGGCGCTTGCAGAGACGGGAGCGGCTTGCAGACTCCCCCAACCAAGAGAGGAGCCCTGAGTC	7680
Dp	7611	CGGCGCTTGCAGAGAGCGGGAGCGGCTTGCAGACTCCCCCAACCAAGAGAGGAGCCCTGAGTC	7670
Qy	7681	CGCGCTGCGCTCATTCCATCTGTCCGTGCAGAGCCGCGATCTTGGCTGTAAAGCCTT	7740
Dp	7671	CGCGCTGCGCTCATTCCATCTGTCCGTGCAGAGCCGCGATCTTGGCTGTAAAGCCTT	7730
Qy	7741	AACCTAAGACTCCCGCCCGCGGCTGGCCCTGTGCAAGCTTTACTAGAGGGAATGTTTACTG	7800
Dp	7731	AACCTAAGACTCCCGCCCGCGGCTGGCCCTGTGCAAGCTTTACTAGAGGGAATGTTTACTG	7790
Qy	7801	GTGCTCGGGAAGGAGAGGGGAGGGGGCGGAGGAGGGGGGACGGACAGGCGTGTGGAGCGAC	7860
Dp	7791	GTGCTCGGGAAGGAGAGGGGAGGGGGCGGAGGAGGGGGGACGGACAGGCGTGTGGAGCGAC	7850
Qy	7861	ACACAGGCGAGCAGAGCGGCGCAGGAGCCCAAGCAGAGTGAACAACGCACTTCAAGCGCAC	7920
Dp	7851	ACACAGGCGAGCAGAGCGGCGCAGGAGCCCAAGCAGAGTGAACAACGCACTTCAAGCGCAC	7910
Qy	7921	TGCGTCCCGGAATGCAATTTGGAACCAAAAGTCTTAACTGAGCTGCGAGCTCCCGCGCCT	7980
Dp	7911	TGCGTCCCGGAATGCAATTTGGAACCAAAAGTCTTAACTGAGCTGCGAGCTCCCGCGCCT	7970

Oy	7981	CCCTCCCGCTCCCAATCCCGCTTACCGCTTCGGAGACAGATGACGGAGCGGAGCCCTGTCAAGCC	8040
Db	7971	CCCTCCCGCTCCCAATCCCGCTTACCGCTTCGGAGACAGATGACGGAGCGGAGCCCTGTCAAGCC	8030
Oy	8041	CCAGTGGCGTCTGTTCGCGGTCCCAACAGCTGCCAGCCAAACGATTTGCTGGAAACCA	8100
Db	8031	CCAATGGCGCTGTTCGCGGTCCCAACAGCTGCCAGCCAAACGATTTGCTGGAAACCA	8090
Oy	8101	GTCAGGCCAAGTGGGCGGACAAAGGCGCAGGTGCGCTTGCGGCGGAAACGAGTGTCCGA	8160
Db	8091	GTCAGGCCAAGTGGGCGGACAAAGGCGCAGGTGCGCTTGCGGCGGAAACGAGTGTCCGA	8150
Oy	8161	GGACTGGAATGTTTTTTTTCACATCGTTCGCGAGCGGTGGGAAAGGAAAGCGAGATGA	8220
Db	8151	GGACTGGAATGTTTTTTTTCACATCGTTCGCGAGCGGTGGGAAAGGAAAGCGAGATGA	8210
Oy	8221	AATGATGATGTGTTTACAGGGGTATATTTTTTGATACCTTCATGATTAATTACAGATGT	8280
Db	8211	AATGATGATGTGTTTACAGGGGTATATTTTTTGATACCTTCATGATTAATTACAGATGT	8270
Oy	8281	TTACGCAAGGAAGACTTACCAGATTAATCTGTGCTGTTTGAATCTCTGCTTACCG	8340
Db	8271	TTACGCAAGGAAGACTTACCAGATTAATCTGTGCTGTTTGAATCTCTGCTTACCG	8330
Oy	8341	TTCAAGAGGCGGTGCGACGCGCAACAGTCGCGGACCCCATCATCTGCAGAACCAAGGGGCG	8400
Db	8331	TTCAAGAGGCGGTGCGACGCGCAACAGTCGCGGACCCCATCATCTGCAGAACCAAGGGGCG	8390
Oy	8401	GGGGAATGCTCGTACAGCCCGCGTGTGCTCCCTCCCTCCCTTCCTTGTGGCAGATGAA	8460
Db	8391	GGGGAATGCTCGTACAGCCCGCGTGTGCTCCCTCCCTCCCTTCCTTGTGGCAGATGAA	8450
Oy	8461	TTTCGATCGATTTCTGTGGCCGCGCATTTTGGCAGAGGTGTGTGTAATTTCTGTCAATTACCA	8520
Db	8451	TTTCGATCGATTTCTGTGGCCGCGCATTTTGGCAGAGGTGTGTGTAATTTCTGTCAATTACCA	8510
Oy	8521	CGTGTCTCTAATTTAAAGGGAATTAATCTCCAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTT	8571
Db	8511	CGTGTCTCTAATTTAAAGGGAATTAATCTCCAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTT	8561

```

RESULT 2
US-11-052-344-1
: Sequence 1, Application US/11052344
: Publication No. US20050191674A1
: GENERAL INFORMATION:
: APPLICANT: Colnec, Lionel
: TITLE OF INVENTION: Method Of Prognosis Of Metastasis By Detection Of FRA12E Fragile Site
: TITLE OF INVENTION: Within The SMRT Gene/Locus At Chromosome 12q24
: FILE REFERENCE: 03551.0181
: CURRENT APPLICATION NUMBER: US/11/052,344
: CURRENT FILING DATE: 2005-02-07
: PRIOR APPLICATION NUMBER: US 60/542,538
: PRIOR FILING DATE: 2004-02-06
: NUMBER OF SEQ ID NOS: 1
: SEQ ID NO 1
: LENGTH: 8561
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: SMRT gene which includes the open reading frame for SMRT protein.
: OTHER INFORMATION: The open reading frame starts at nucleotide 2 and ends at nucleotide
US-11-052-344-1

```

Query Match Similarity 99.7%; Score 8541; DB 10; Length 8561;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 8560; Conservative 1; Mismatches 0; Indels 10; Gaps 1;

QY 61 GCCCAGAGCCTTTCTTACCCAGTGCAGATCGCCCGAGCGCAACGAGAGTTCGGGCTCCT 120  
Db 61 GCCCAGAGCCTTTCTTACCCAGTGCAGATCGCCCGAGCGCAACGAGAGTTCGGGCTCCT 120  
QY 121 GAGAGTACAGAGCACTCCCGGACTATATGCTTCCCACTGTGCGGGGTCCATCATCCA 180  
Db 121 GAGAGTACAGAGCACTCCCGGACTATATGCTTCCCACTGTGCGGGGTCCATCATCCA 180  
QY 181 GCCCCAGGCGGAGGAGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGAAATGAACGGTCCCA 240  
Db 181 GCCCCAGGCGGAGGAGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGAAATGAACGGTCCCA 240  
QY 241 GAGAGTCCACCTGCGGCGCAAGTCCCATCATATCCTGCCGAGCTGGGGAATCAGAGAT 300  
Db 241 GAGAGTCCACCTGCGGCGCAAGTCCCATCATATCCTGCCGAGCTGGGGAATCAGAGAT 300  
QY 301 GAGATTCAATTGAAAGCAAGCCCTTCGGCTTAAAGCTGCTGCCTGACCCCTTCTCTGAGCC 360  
Db 301 GAGATTCAATTGAAAGCAAGCCCTTCGGCTTAAAGCTGCTGCCTGACCCCTTCTCTGAGCC 360  
QY 361 GTCAACCCCTGCTGGCGACGGGGCAGGCTGCGGGATCTGAAAGACTCAACCAAGAACCGTAG 420  
Db 361 GTCAACCCCTGCTGGCGACGGGGCAGGCTGCGGGATCTGAAAGACTCAACCAAGAACCGTAG 420  
QY 421 CCTGAACGGGCAAGCTGGAACCGGTGTCTTCCCGCAGCCCGCCGCAACTGAACCTGAGCT 480  
Db 421 CCTGAACGGGCAAGCTGGAACCGGTGTCTTCCCGCAGCCCGCCGCAACTGAACCTGAGCT 480  
QY 481 GAGAGTGTGCTGCGCAACGGCTGTCCAAGAGAGAGCTGATCCGAACATGGAACCGGTGGA 540  
Db 481 GAGAGTGTGCTGCGCAACGGCTGTCCAAGAGAGAGCTGATCCGAACATGGAACCGGTGGA 540  
QY 541 CCGAGAGATCAACATGATGATGAGCAGACAGATCTCTAAGCTGAAGAAAGAACAGCAAGCT 600  
Db 541 CCGAGAGATCAACATGATGATGAGCAGACAGATCTCTAAGCTGAAGAAAGAACAGCAAGCT 600  
QY 601 GAGAGAGAGAGGCTGCGCAAGCCCGCCGAGCTGAGAAAGCCCGGTGTCAACCGCCGATCGA 660  
Db 601 GAGAGAGAGAGGCTGCGCAAGCCCGCCGAGCTGAGAAAGCCCGGTGTCAACCGCCGATCGA 660  
QY 661 GTCGAAGCAACCGGACCTGTGTGCAAGTCAATCTAAGCGAAGAACCCGAGAGAGGCTGAAAGC 720  
Db 661 GTCGAAGCAACCGGACCTGTGTGCAAGTCAATCTAAGCGAAGAACCCGAGAGAGGCTGAAAGC 720  
QY 721 TGCAATGCGGAATTTCTGGAAGGCTGCGGGGCGCCAGGCTGAGAGCTGCGGCTGTACAACAGGC 780  
Db 721 TGCAATGCGGAATTTCTGGAAGGCTGCGGGGCGCCAGGCTGAGAGCTGCGGCTGTACAACAGGC 780  
QY 781 CTCCGACACCCCGGCAAGTATCATGAGAACATCAAAATTAACAGGCGATGCGGAAGAGCT 840  
Db 781 CTCCGACACCCCGGCAAGTATCATGAGAACATCAAAATTAACAGGCGATGCGGAAGAGCT 840  
QY 841 AATCTTTGTAATTCAAGAGAGAGATCAAGCTTGGAAACAATGGAAGCAGAAATTTCTGCCA 900  
Db 841 AATCTTTGTAATTCAAGAGAGAGATCAAGCTTGGAAACAATGGAAGCAGAAATTTCTGCCA 900  
QY 901 GGGGCTATGACCAAGCTCATGGAAGGCTTGGAAAAAAAGGGAAGCGGATGAAAAACAACC 960  
Db 901 GGGGCTATGACCAAGCTCATGGAAGGCTTGGAAAAAAAGGGAAGCGGATGAAAAACAACC 960  
QY 961 GGGCGGCGGCGCCAGAGAGAGCAAGGTGCGCGAGTACTACGAAAAAGCAATTTCCCTGAGAT 1020  
Db 961 GGGCGGCGGCGCCAGAGAGAGCAAGGTGCGCGAGTACTACGAAAAAGCAATTTCCCTGAGAT 1020  
QY 1021 CCGCAAGCAACCGGAGCTGCAAGAGCGCATGCAAGACAGGAGTGGGCGAGCGGCGGCAAGTGG 1080  
Db 1021 CCGCAAGCAACCGGAGCTGCAAGAGCGCATGCAAGACAGGAGTGGGCGAGCGGCGGCAAGTGG 1080  
QY 1081 GGTGTTCATGTGGCGCGCCGAGAGAGCAAGAGGTTCAGAGATTCATGATGGCCCTCTC 1140  
Db 1081 GGTGTTCATGTGGCGCGCCGAGAGAGCAAGAGGTTCAGAGATTCATGATGGCCCTCTC 1140  
QY 1141 AGAGCAGAGAACTTGAGAAAGCAAGATGCGCAGCTGGCGGTGATCCCGCATGCTGTA 1200

Db 1141 AGAGCAGAGAACTTGAGAAAGCAAGATGCGCAGCTGGCGGTGATCCCGCATGCTGTA 1200  
QY 1201 CGAGCTGACCAAGCAGGACATCAAGTTTCATCAACAATGAACGGGCTTATGCGCGAACCCAT 1260  
Db 1201 CGAGCTGACCAAGCAGGACATCAAGTTTCATCAACAATGAACGGGCTTATGCGCGAACCCAT 1260  
QY 1261 GAAGGTGTACAAAGACCGCAGCTCATGACAATGTGAGTGAAGCAGAGAAAGAGACCTT 1320  
Db 1261 GAAGGTGTACAAAGACCGCAGCTCATGACAATGTGAGTGAAGCAGAGAAAGAGACCTT 1320  
QY 1321 CCGGAGAAAGTTTCATGCAAGATCCCAAGAACTTTGGCTGATCCGATATTTCTTGAAGAG 1380  
Db 1321 CCGGAGAAAGTTTCATGCAAGATCCCAAGAACTTTGGCTGATCCGATATTTCTTGAAGAG 1380  
QY 1381 GAAGCAGTGGCTGATGTCGCTCTTATTAATTAACCTGACTAAGAAATGAGAACTTATTA 1440  
Db 1381 GAAGCAGTGGCTGATGTCGCTCTTATTAATTAACCTGACTAAGAAATGAGAACTTATTA 1440  
QY 1441 GAGCTGTGTGAGACGAGACTATTCGCGCGCGGCAAGAGCCAGCAACACAGACGA 1500  
Db 1441 GAGCTGTGTGAGACGAGACTATTCGCGCGCGGCAAGAGCCAGCAACACAGACGA 1500  
QY 1501 GCAGCAGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1560  
Db 1501 GCAGCAGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1560  
QY 1561 AGATGAGAAAGAGAAAGAAAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
Db 1561 AGATGAGAAAGAGAAAGAAAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
QY 1621 CGACAAAGAGAGACCTCTCAAGAGAGAGACAGACGACCTCAAGGAGAGAGACAAAGACGA 1680  
Db 1621 CGACAAAGAGAGACCTCTCAAGAGAGAGAGACAGACGACCTCAAGGAGAGAGACAAAGACGA 1680  
QY 1681 GAAGAGGCTGTGCTTCCAAAAGGCTCGCAAACTTGCACACGCCAGGAGAGACGCAAAAG 1740  
Db 1681 GAAGAGGCTGTGCTTCCAAAAGGCTCGCAAACTTGCACACGCCAGGAGAGACGCAAAAG 1740  
QY 1741 CCGCATCAACCGGCTCAATGCTAATGAGGCCCAAGAGAGAGGCTCATCCCCCAGCA 1800  
Db 1741 CCGCATCAACCGGCTCAATGCTAATGAGGCCCAAGAGAGAGGCTCATCCCCCAGCA 1800  
QY 1801 GAGCGCCGAGGCTGCTCATGAGAGCTGAATGAGAGTTCTGCTGGAAGAGAGAGAAAT 1860  
Db 1801 GAGCGCCGAGGCTGCTCATGAGAGCTGAATGAGAGTTCTGCTGGAAGAGAGAGAAAT 1860  
QY 1861 GGAACACGCCCAAGAAAGGTCTCTGGAACAGCGCGCAACTGTGTGCGCCATGCGCCGAT 1920  
Db 1861 GGAACACGCCCAAGAAAGGTCTCTGGAACAGCGCGCAACTGTGTGCGCCATGCGCCGAT 1920  
QY 1921 GGTGGGCTCCAAAGCTGTGTGCGCAGTGTAAGAACTTCACTTCAACTCAAGAAAGGCA 1980  
Db 1921 GGTGGGCTCCAAAGCTGTGTGCGCAGTGTAAGAACTTCACTTCAACTCAAGAAAGGCA 1980  
QY 1981 GAACCTCGATGAGATCTTGCAGACAGCAAGCTGAAATGAGTGAAGAGAGAGAAAGCGCG 2040  
Db 1981 GAACCTCGATGAGATCTTGCAGACAGCAAGCTGAAATGAGTGAAGAGAGAGAAAGCGCG 2040  
QY 2041 GAGAAAGAAAGAAAGAGCGCGCGCGCGCGCGCGCGCGAGAGAGCTGCAATTCGCGCCGTGCT 2100  
Db 2041 GAGAAAGAAAGAAAGAGCGCGCGCGCGCGCGCGCGCGCGAGAGAGCTGCAATTCGCGCCGTGCT 2100  
QY 2101 GAGAGATGAGAGAGATGAGAGGCTGCGGGGTGAAGCGGAAATGAGAGAGATGTGAGAGGA 2160  
Db 2101 GAGAGATGAGAGAGATGAGAGGCTGCGGGGTGAAGCGGAAATGAGAGAGATGTGAGAGGA 2160  
QY 2161 GAGTGAAGCTTTAATGAGCTCTGGAAGTGAAGTGGCCAGAGGAGAAATGAGAGGCGCCAGC 2220  
Db 2161 GAGTGAAGCTTTAATGAGCTCTGGAAGTGAAGTGGCCAGAGGAGAAATGAGAGGCGCCAGC 2220  
QY 2221 CACTGTCAACAAACAGCTCAGACCCGAGAGCAATCCCTCTCTCTCACTGAGGCGCCAA 2280



Db 2221 CACTGTCAACACAGCTCAGACACCGAGAGATCCCCCTCTCTCACATGAGGCCGCCAA 2280  
QY 2281 GACACAGGGCAGAAATGGGCCCCAAGCCCCCGACGACCTTGGGGCGGACGGGCAACCCC 2240  
Db 2281 GACACAGGGCAGAAATGGGCCCCAAGCCCCCGACGACCTTGGGGCGGACGGGCAACCCC 2240  
QY 2341 AGGCCCAACCCCAACCAACGAGAGCATCCCGGGCCCCCATTTGAGCCCAACCCCGGCTTC 2400  
Db 2341 AGGCCCAACCCCAACCAACGAGAGCATCCCGGGCCCCCATTTGAGCCCAACCCCGGCTTC 2400  
QY 2401 TGAAGGACACCGAGGCCCTTACGCCCCCAACGACATCCCTCATGCTCTGCACTCTCTCC 2460  
Db 2401 TGAAGGACACCGAGGCCCTTACGCCCCCAACGACATCCCTCATGCTCTGCACTCTCTCC 2460  
QY 2461 TGTGGTCCCCAAG 2520  
Db 2461 TGTGGTCCCCAAG 2520  
QY 2521 GAG 2580  
Db 2521 GAG 2580  
QY 2581 GCCCGCTCAAG 2640  
Db 2581 GCCCGCTCAAG 2640  
QY 2641 GAG 2700  
Db 2641 GAG 2700  
QY 2701 CGGAG 2760  
Db 2701 CGGAG 2760  
QY 2761 CTGCAAGTGAAG 2820  
Db 2761 CTGCAAGTGAAG 2820  
QY 2821 AAGGCCAGAGCTCTCAACCCCACTGGGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
Db 2821 AAGGCCAGAGCTCTCAACCCCACTGGGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
QY 2881 ACTGAGCTGAG 2940  
Db 2881 ACTGAGCTGAG 2940  
QY 2941 AGTCCATGAG 3000  
Db 2941 AGTCCATGAG 3000  
QY 3001 GGCACCGCAAAACCTGCAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
Db 3001 GGCACCGCAAAACCTGCAACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
QY 3061 GGGCAAG 3120  
Db 3061 GGGCAAG 3120  
QY 3121 GAAAGCTGCTGGGAG 3180  
Db 3121 GAAAGCTGCTGGGAG 3180  
QY 3181 TGAAGTGAATCAAG 3240  
Db 3181 TGAAGTGAATCAAG 3240  
QY 3241 TGGTCAACCACTGCGCTGGGCTTCAATGACATGCGCGGCGGCTCTGCGGCGGCAAC 3300  
Db 3241 TGGTCAACCACTGCGCTGGGCTTCAATGACATGCGCGGCGGCTCTGCGGCGGCAAC 3300  
QY 3301 CACCATCTCAACCGGCT 3360  
Db 3301 CACCATCTCAACCGGCT 3360

QY 3361 GCAAATAGGTGACCATCTCCCAAGGAATGTGGGTCAAGTTCACAGTCCCGTACTCAGAGCA 3420  
Db 3361 GCAAATAGGTGACCATCTCCCAAGGAATGTGGGTCAAGTTCACAGTCCCGTACTCAGAGCA 3420  
QY 3421 TGCCAAGGCGCGGTGGGCGCTGTCAACATGGGGCTGCGCTTCCCATTTGAGACCCCAAAA 3480  
Db 3421 TGCCAAGGCGCGGTGGGCGCTGTCAACATGGGGCTGCGCTTCCCATTTGAGACCCCAAAA 3480  
QY 3481 GCTGGGACCTTCAAGGAG 3540  
Db 3481 GCTGGGACCTTCAAGGAG 3540  
QY 3541 ACCGAG 3600  
Db 3541 ACCGAG 3600  
QY 3601 GGGCTCAGTTCCGGGCGGAAAGCATCAAAAGGCAATTCACAGCAACGGGTGCTCTCGGA 3660  
Db 3601 GGGCTCAGTTCCGGGCGGAAAGCATCAAAAGGCAATTCACAGCAACGGGTGCTCTCGGA 3660  
QY 3661 CAGCGGCATCAATACCGGGGCTTCATCAACCGGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 3720  
Db 3661 CAGCGGCATCAATACCGGGGCTTCATCAACCGGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 3720  
QY 3721 GGGCAACATCAACAGGATCATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780  
Db 3721 GGGCAACATCAACAGGATCATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780  
QY 3781 GGAAGAGCTGCCCAAG 3840  
Db 3781 GGAAGAGCTGCCCAAG 3840  
QY 3841 TGAAGGTGGATGTCTGTGATCCAGTGTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3900  
Db 3841 TGAAGGTGGATGTCTGTGATCCAGTGTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3900  
QY 3901 CCCCCATGAG 3960  
Db 3901 CCCCCATGAG 3960  
QY 3961 AGGCATCTCTCAAG 4020  
Db 3961 AGGCATCTCTCAAG 4020  
QY 4021 CAGCCCCCAACCTTAAG 4080  
Db 4021 CAGCCCCCAACCTTAAG 4080  
QY 4081 TCGGTCCTACGTGAG 4140  
Db 4081 TCGGTCCTACGTGAG 4140  
QY 4141 GAGAGGACAGCTCTCGCCCCCAACCGGCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4200  
Db 4141 GAGAGGACAGCTCTCGCCCCCAACCGGCTTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4200  
QY 4201 GGGCCTGGGCGGCTTGAAGCTGAAGCGGCGCATGAGGGGCTGTGTGAGAGAGAGAGAGAGAG 4260  
Db 4201 GGGCCTGGGCGGCTTGAAGCTGAAGCGGCGCATGAGGGGCTGTGTGAGAGAGAGAGAGAGAG 4260  
QY 4261 GGGGGGGGCTCATATCAGAGATCCCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4320  
Db 4261 GGGGGGGGCTCATATCAGAGATCCCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4320  
QY 4321 CCTGGCCCCGCGCGCTTCAAG 4380  
Db 4321 CCTGGCCCCGCGCGCTTCAAG 4380  
QY 4381 CACCGGCGGCTCAACATCTGCTTCCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4440  
Db 4381 CACCGGCGGCTCAACATCTGCTTCCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4440

QY	4441	CGGCGGAGCGTTCCACCCCGTGACACCGCTGAGATGTATATGGCCGACGCCCGCGGACCTGGA	4500
Db	4441	CGGCGGAGCGTTCCACCCCGTGACACCGCTGAGATGTATATGGCCGACGCCCGCGGACCTGGA	4500
QY	4501	ACGTCCTGCTACGAGAGAGACCTTGAAAGCGCGGCAGAGGACCCGCACAGCTCGAGGAGG	4560
Db	4501	ACGTCCTGCTACGAGAGAGACCTTGAAAGCGCGGCAGAGGACCCGCACAGCTCGAGGAGG	4560
QY	4561	CTCCATTGCGCGCGCGCCCGCGTCAATTGTGCTTGAAGTGGTAAGCTCGCGGACGAGCC	4620
Db	4561	CTCCATTGCGCGCGCGCCCGCGTCAATTGTGCTTGAAGTGGTAAGCTCGCGGACGAGCC	4620
QY	4621	CCTGACCTATAGAGACCAACGGGGACACCTTTGGCGGCACTCTCCAGAGAGTTGGCCCT	4680
Db	4621	CCTGACCTATAGAGACCAACGGGGACACCTTTGGCGGCACTCTCCAGAGAGTTGGCCCT	4680
QY	4681	GACCATGCGGGAGGCCACGCGCGCTTGACAGAGAGGAGCCTTTCGTTCAGACAGGCATC	4740
Db	4681	GACCATGCGGGAGGCCACGCGCGCTTGACAGAGAGGAGCCTTTCGTTCAGACAGGCATC	4740
QY	4741	CCAGGACCGAAAAGCTGACGCTGACGCGCTTGTAAGATGCGCAAGTCCCGCACAGACCGT	4800
Db	4741	CCAGGACCGAAAAGCTGACGCTGACGCGCTTGTAAGATGCGCAAGTCCCGCACAGACCGT	4800
QY	4801	GCCCGAGACCAACCCACACCCCATTCGCGCTTATGAGACACTGCTTGGGGCGTGAATGG	4860
Db	4801	GCCCGAGACCAACCCACACCCCATTCGCGCTTATGAGACACTGCTTGGGGCGTGAATGG	4860
QY	4861	CGTGACCTGATGCGAGCGACATCCCGCTGAGCTTGACCCCACTTCATACCCCGCG	4920
Db	4861	CGTGACCTGATGCGAGCGACATCCCGCTGAGCTTGACCCCACTTCATACCCCGCG	4920
QY	4921	CATCCCTGTGACGCGACGCGCTGCTACTACCTGCGCCGACACCTGGCGCCCAACCCAC	4980
Db	4921	CATCCCTGTGACGCGACGCGCTGCTACTACCTGCGCCGACACCTGGCGCCCAACCCAC	4980
QY	4981	CTACCCGCACTGTATCCCACTTATCCGAGCTTATCCCGAGCTTACCCGAGCGCTGGA	5040
Db	4981	CTACCCGCACTGTATCCCACTTATCCGAGCTTATCCCGAGCTTATCCCGAGCGCTGGA	5040
QY	5041	GAAACCGGAGACCATCATCATATGATATATACCTCGAGAGATGACCAACAAGGC	5100
Db	5041	GAAACCGGAGACCATCATCATATGATATATACCTCGAGAGATGACCAACAAGGC	5100
QY	5101	CACCGCACATGACCCGAGCGCTGATATCTGAGGGGCTCTCGACCGCGAGTCTCGCT	5160
Db	5101	CACCGCACATGACCCGAGCGCTGATATCTGAGGGGCTCTCGACCGCGAGTCTCGCT	5160
QY	5161	GGCACTCAACTACGCTGCGGGTCCCGAGGCATCATGACTGTGCCAAGTGCACACT	5220
Db	5161	GGCACTCAACTACGCTGCGGGTCCCGAGGCATCATGACTGTGCCAAGTGCACACT	5220
QY	5221	GCTGTGCTCGTCCCCCGGACACCAAGGACCCGACCAACGCGCATGAGCGCTTGGCTA	5280
Db	5221	GCTGTGCTCGTCCCCCGGACACCAAGGACCCGACCAACGCGCATGAGCGCTTGGCTA	5280
QY	5281	CCTCCGACCGGCGCCCGACCCCTTGACGAGCGGCAACGAGCTCCCCACTCTCCGAGG	5340
Db	5281	CCTCCGACCGGCGCCCGACCCCTTGACGAGCGGCAACGAGCTCCCCACTCTCCGAGG	5340
QY	5341	AGGTCCAACACACTTGAACAAACCAACCAACGCTCTGTCGAGCGGAGCGAGACCG	5400
Db	5341	AGGTCCAACACACTTGAACAAACCAACCAACGCTCTGTCGAGCGGAGCGAGACCG	5400
QY	5401	GGATTGAGAGCGGGACCGGATTCGGAGACGGGAAAAGTCACTCTCAAGTCCACACAGAC	5460
Db	5401	GGATTGAGAGCGGGACCGGATTCGGAGACGGGAAAAGTCACTCTCAAGTCCACACAGAC	5460
QY	5461	GGTGAGACGCAACCATCTGAGACCTGATACAGAGACGAGCGGACGAGCGGAG	5520
Db	5461	GGTGAGACGCAACCATCTGAGACCTGATACAGAGACGAGCGGACGAGCGGAG	5520
QY	5521	CAGCGCGCGGGGTGGGGCAGACGACGCGCGCTCCCACTCCATGCGCCACGCA	5580

Db	5521	CAGCGCGCGGGGTGGGGGACGAGAGCGCGCCCGCTCCACTCCCAATGCCACACGCA	5580
QY	5581	CTCGCCCATCTCCCTCGGACCCAGAGATGCCCTCCAGACAGAACCCAGTGTGTTCAAA	5640
Db	5581	CTCGCCCATCTCCCTCGGACCCAGAGATGCCCTCCAGACAGAGACCCAGTGTGTTCAAA	5640
QY	5641	CACAGGCAATGAAGGGTATCATACCGCGTGTGGAGGCCACGAAGCCCAAGCTCTTGAGGTC	5700
Db	5641	CACAGGCAATGAAGGGTATCATACCGCGTGTGGAGGCCACGAAGCCCAAGCTCTTGAGGTC	5700
QY	5701	CACCTTCACCTCTCAACCGGTTTGGCCCACTCAGACATTTCCACTTGGCCACCACTTGGCC	5760
Db	5701	CACCTTCACCTCTCAACCGGTTTGGCCCACTCAGACATTTCCACTTGGCCACCACTTGGCC	5760
QY	5761	ACTGGGCGGACCCCTCGATGGGGTCTACCTTACCTCATGAGGCCCGCTTGTGTCGCAA	5820
Db	5761	ACTGGGCGGACCCCTCGATGGGGTCTACCTTACCTCATGAGGCCCGCTTGTGTCGCAA	5820
QY	5821	GAAGGCCCCCGGGGTGCGCCGGCCAGAGGGGCCCGGAGCAGACATCGGCATATGCTTTCT	5880
Db	5821	GAAGGCCCCCGGGGTGCGCCGGCCAGAGGGGCCCGGAGCAGACATCGGCATATGCTTTCT	5880
QY	5881	CGCCAGGCCCCCAAGCCCGGTCCGGGCTGGAGCCCGCTTCTCCCAAGAGGGCTTCGGA	5940
Db	5881	CGCCAGGCCCCCAAGCCCGGTCCGGGCTGGAGCCCGCTTCTCCCAAGAGGGCTTCGGA	5940
QY	5941	GCCCCGCGCCCTTAGTGCCTCTGTCTTGGCCACGCAACATGACCAGACCCCTGCGAA	6000
Db	5941	GCCCCGCGCCCTTAGTGCCTCTGTCTTGGCCACGCAACATGACCAGACCCCTGCGAA	6000
QY	6001	GAACCTCGCATCTCAACACGCCAGCCCGGACCCCGCGGCGCACTGTGCTCGGCTTCGGA	6060
Db	6001	GAACCTCGCATCTCAACACGCCAGCCCGGACCCCGCGGCGCACTGTGCTCGGCTTCGGA	6060
QY	6061	CCCGGACCGGGAAGAAGTCAATGAATTAACCTTTTCATTCGGAATCTGAAATCTCGTTTC	6120
Db	6061	CCCGGACCGGGAAGAAGTCAATGAATTAACCTTTTCATTCGGAATCTGAAATCTCGTTTC	6120
QY	6121	TCTGGGTTTACCAACGCGACGACTAAGCCCCGGAAGGGGTGAGGCCGCTGTGAG	6180
Db	6121	TCTGGGTTTACCAACGCGACGACTAAGCCCCGGAAGGGGTGAGGCCGCTGTGAG	6180
QY	6181	CTCACCCAGTCTGACCCACGACGACAGAGGGGCTCCCAAGCACTTGAAGAGCTGCACAAAG	6240
Db	6181	CTCACCCAGTCTGACCCACGACGACAGAGGGGCTCCCAAGCACTTGAAGAGCTGCACAAAG	6240
QY	6241	CCACTTGAAGGGGGAGCTGCGGCCCAAGAGCAGGCCCCCGTGAACCTTGGCGGGAGGC	6300
Db	6241	CCACTTGAAGGGGGAGCTGCGGCCCAAGAGCAGGCCCCCGTGAACCTTGGCGGGAGGC	6300
QY	6301	CGCCACCTCCCAACCTGCGGCGCTGCTGAGAGCAAGCCCTCGTCCAGCCCGCTGCT	6360
Db	6301	CGCCACCTCCCAACCTGCGGCGCTGCTGAGAGCAAGCCCTCGTCCAGCCCGCTGCT	6360
QY	6361	CCAGAACCGCCCAAGGGGTCAAAAGTTCACCAAGCGAGTGTTCACCTTGGCCACCATCAG	6420
Db	6361	CCAGAACCGCCCAAGGGGTCAAAAGTTCACCAAGCGAGTGTTCACCTTGGCCACCATCAG	6420
QY	6421	TGAGGTCAATCACAGAGATCACCCCGGACCAACCCACAGCACTCAGGCAACCTTGGC	6480
Db	6421	TGAGGTCAATCACAGAGATCACCCCGGACCAACCCACAGCACTCAGGCAACCTTGGC	6480
QY	6481	CGCCCCCTCTAATCTTCTTCCCTGAGGGCCAGCTGCCCCGTCTTGAACCTTCGCGGCCAC	6540
Db	6481	CGCCCCCTCTAATCTTCTTCCCTGAGGGCCAGCTGCCCCGTCTTGAACCTTCGCGGCCAC	6540
QY	6541	CAGTGAACCTTACCTTCCGCGCCCGGACCATATGTGCCCCGGCCCGTGGCTCCCCCAG	6600
Db	6541	CAGTGAACCTTACCTTCCGCGCCCGGACCATATGTGCCCCGGCCCGTGGCTCCCCCAG	6600
QY	6601	CGAAGGGGACAGAGGTCTTCCAGAGCCAAACAGACGTGCTTGGGTGTGTGAGGA	6660

6601 CGAAGGGGCAAGAGTCTCCAGAGCCAAACAGAGTCTGTTGGTGTGAGGA 6660  
6661 CGGTATTGAACCTGTGTCCCAACGGAGGGCAATGACGAGCCAGGCACTCCCGAGTGC 6720  
6661 CGGTATTGAACCTGTGTCCCAACGGAGGGCAATGACGAGCCAGGCACTCCCGAGTGC 6720  
6721 TGTGTATCCCGTGTGTATCCGGGATGGGAAACAGAGCCAGAGGATGGGCTTCAA 6780  
6721 TGTGTATCCCGTGTGTATCCGGGATGGGAAACAGAGCCAGAGGATGGGCTTCAA 6780  
6781 GTTCTCAGGCAACACCAAGCCAGCCGCTTCTTCAAGAGCTGACCGAGCAATC 6840  
6781 GTTCTCAGGCAACACCAAGCCAGCCGCTTCTTCAAGAGCTGACCGAGCAATC 6840  
6841 CGGCAATGTCTCAAGTCCAAAGAGAGATCAACAGAGCTGAAACCCCAACCGGAA 6900  
6841 CGGCAATGTCTCAAGTCCAAAGAGAGATCAACAGAGCTGAAACCCCAACCGGAA 6900  
6901 TGAGCTTGAAATCAATATGAGCCAGGCTGGGACGAGATCTTCAATATGCTCCGATCAC 6960  
6901 TGAGCTTGAAATCAATATGAGCCAGGCTGGGACGAGATCTTCAATATGCTCCGATCAC 6960  
6961 CGGAAACAGGCTTTATGACCTTATGAAAGCAAGCGGTGCAAGAACATGCAACAT 7020  
6961 CGGAAACAGGCTTTATGACCTTATGAAAGCAAGCGGTGCAAGAACATGCAACAT 7020  
7021 GGGGCTGAGGGCAATTAAGAAAGGCACTCAGTGGTAAATATGACAGTGGGAAAGTCT 7080  
7021 GGGGCTGAGGGCAATTAAGAAAGGCACTCAGTGGTAAATATGACAGTGGGAAAGTCT 7080  
7081 CCCGCGCTCAGCGCAATGCTTTTAACTCTGAAATGCGAGTGCAGGCTGCGCTGC 7140  
7081 CCCGCGCTCAGCGCAATGCTTTTAACTCTGAAATGCGAGTGCAGGCTGCGCTGC 7140  
7141 TATGCCATTAACCGCTGTGACGAGCGAGTGAACAACAATCACTCTGCGCAGTGGCG 7200  
7141 TATGCCATTAACCGCTGTGACGAGCGAGTGAACAACAATCACTCTGCGCAGTGGCG 7200  
7201 CGGGAAGGCAAGGTCTTGGGAGAGCCAGCAAGCCCAAGCAATCCCGGCGCGCG 7260  
7201 CGGGAAGGCAAGGTCTTGGGAGAGCCAGCAAGCCCAAGCAATCCCGGCGCGCG 7260  
7261 CTTGGCATCTGGGAGACCGGCTCCTCTGTCTCTCAATGAGCTGCGAGGAGATCGCA 7320  
7261 CTTGGCATCTGGGAGACCGGCTCCTCTGTCTCTCAATGAGCTGCGAGGAGATCGCA 7320  
7321 CGGCGGAGCGGCTGACCAACCGGCTGTGGAGAGACAGGCGCTGCTCGCAGTTCAC 7380  
7321 CGGCGGAGCGGCTGACCAACCGGCTGTGGAGAGACAGGCGCTGCTCGCAGTTCAC 7380  
7381 GGCATTCCCTTCAACACCCCTGATCATGCGGCTGCGAGGCGGTGTATGGCTTCCAC 7440  
7381 GGCATTCCCTTCAACACCCCTGATCATGCGGCTGCGAGGCGGTGTATGGCTTCCAC 7440  
7441 CCCACCGGCTCTCCCGCGGAGCGAGGCGGCTCTGCTGAGGCGGCAACAGGCTGAGCA 7500  
7441 CCCACCGGCTCTCCCGCGGAGCGAGGCGGCTCTGCTGAGGCGGCAACAGGCTGAGCA 7500  
7501 GGAAGCCCAAGCACTGCTGTCTGCTGCAAGTACAGACACTTCCGACAGGAGTACTCAG 7560  
7501 GGAAGCCCAAGCACTGCTGTCTGCTGCAAGTACAGACACTTCCGACAGGAGTACTCAG 7560  
7561 AACAGGGCGGG 7610  
7561 AACAGGGCGGG 7610  
7621 CGGAGCTGAGAGCGGGGCGGCTGCGCACTCCCAACCAAGAGAGAGGCGGCTGAGTCT 7680  
7621 CGGAGCTGAGAGCGGGGCGGCTGCGCACTCCCAACCAAGAGAGAGGCGGCTGAGTCT 7680  
7681 CGGCTGAGAGCGGGGCGGCTGCGCACTCCCAACCAAGAGAGAGGCGGCTGAGTCT 7740  
7681 CGGCTGAGAGCGGGGCGGCTGCGCACTCCCAACCAAGAGAGAGGCGGCTGAGTCT 7740  
7671 CGGCTGAGAGCGGGGCGGCTGCGCACTCCCAACCAAGAGAGAGGCGGCTGAGTCT 7730

7741 AACTAAGACTCCCGGCTGCGCTGTCAGACCTTACTCAGGGAGTGTATCTG 7800  
7731 AACTAAGACTCCCGGCTGCGCTGTCAGACCTTACTCAGGGAGTGTATCTG 7790  
7801 GTTCTCAGGCAACACCAAGCCAGCCGCTTCTTCAAGAGCTGACCGAGCAATC 7860  
7791 GTTCTCAGGCAACACCAAGCCAGCCGCTTCTTCAAGAGCTGACCGAGCAATC 7850  
7861 ACACAGGCGGCAAGGGCGGCAAGGGCCCAAGAGAGATGACAGCACTTCCAGCCAC 7920  
7851 ACACAGGCGGCAAGGGCGGCAAGGGCCCAAGAGAGATGACAGCACTTCCAGCCAC 7910  
7921 TGCCTCCCGAATGATTTGGAACCAAGTCTTAACTGAGTCTGAGGCGGCGGCT 7980  
7911 TGCCTCCCGAATGATTTGGAACCAAGTCTTAACTGAGTCTGAGGCGGCGGCT 7970  
7981 CCTCCGCTCCCATCCGCTTACGCTGTGACAGATGAGAGCGAGGCGGCTGTCAGCC 8040  
7971 CCTCCGCTCCCATCCGCTTACGCTGTGACAGATGAGAGCGAGGCGGCTGTCAGCC 8030  
8041 CGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8100  
8031 CGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8090  
8101 GTCAAGGCAAGTGGCGGCAAAAGGCGAGTGGCGGCTGCGGAGGAGAGATGCTCG 8160  
8091 GTCAAGGCAAGTGGCGGCAAAAGGCGAGTGGCGGCTGCGGAGGAGAGATGCTCG 8150  
8161 GGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8220  
8151 GGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8210  
8221 AATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8280  
8211 AATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8270  
8281 TTAAGCAAGAGAGCTTACCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8340  
8271 TTAAGCAAGAGAGCTTACCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8330  
8341 TTAAGCAAGAGAGCTTACCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8400  
8331 TTAAGCAAGAGAGCTTACCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8390  
8401 GGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8460  
8391 GGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8450  
8461 TTAAGCAAGAGAGCTTACCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8520  
8451 TTAAGCAAGAGAGCTTACCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8510  
8521 CGTGTCTTAAATTAAGAGCAATTAATCTCAAAAAA 8571  
8511 CGTGTCTTAAATTAAGAGCAATTAATCTCAAAAAA 8561

RESULT 3  
US-10-087-192-653  
; Sequence 653, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586

;; PRIOR FILING DATE: 2001-03-02  
;; NUMBER OF SEQ ID NOS: 2059  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 653  
;; LENGTH: 8667  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-087-192-653

Query Match 98.8%; Score 8459.2; DB 5; Length 8667;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 8512; Conservative 1; Mismatches 38; Indels 11; Gaps 2;

QY 1 CATGTGCGGCTTCACACACAGCTTTGTGGCACAAGCTGGAGGGCCACTGAGCCCGCTAACCC 60  
DB CATGTGCGGATCCACACAGCTGTGGCACAAGCTGGAGGGCCACTGAGCCCGCTAACCC 174  
QY 61 GCCCCACAGCCTTTCTTACCCAGTGCAGATCGCCCGGACGACAGGAGTGGGCTCTT 120  
DB 175 GCCCCACAGCCTTTCTTACCCAGTGCAGATCGCCCGGACGACAGGAGTGGGCTCTT 234  
QY 121 GGAAGTACAGCAGCACTCCCGGACTATGCTCCACCTGTGCGCGGGCTCCATCATCCA 180  
DB 235 GGAAGTACAGCAGCACTCCCGGACTATGCTCCACCTGTGCGCGGGCTCCATCATCCA 294  
QY 181 GCCCAGCGCGGAGGCGCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCA 240  
DB 295 GCCCAGCGCGGAGGCGCTCCCTGCTGTCTGAGTTCCAGCCCGGGAATGAACGGTCCCA 354  
QY 241 GGAAGTCCACCTCGCGCCAGAGTCCCACTCATCTGCCCCAGCTGGGGAATGCAGAGT 300  
DB 355 GGAAGTCCACCTCGCGCCAGAGTCCCACTCATCTGCCCCAGCTGGGGAATGCAGAGT 414  
QY 301 GGAAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCGCTGACCCCTGCTGGAGCC 360  
DB 415 GGAAGTTCATTGAAAGCAAGCGCCCTCGGCTAGAGCTGCGCTGACCCCTGCTGGAGCC 474  
QY 361 GTCAACCCCTGTGCGCACAGGGCCAGCTTGGAGATCTGAAAGACTTCAACCAAGACCGTAG 420  
DB 475 GTCAACCCCTGTGCGCACAGGGCCAGCTTGGAGATCTGAAAGACTTCAACCAAGACCGTAG 534  
QY 421 CCTGACCGGCAAGCTGGAAACCGGTGTCTTCCCCAGCCCCCGCACACTGACCTTGAGCT 480  
DB 535 CCTGACCGGCAAGCTGGAAACCGGTGTCTTCCCCAGCCCCCGCACACTGACCTTGAGCT 594  
QY 481 GGAAGCTGGTGGCGCACCGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGAACCGGTGGA 540  
DB 595 GGAAGCTGGTGGCGCACCGGCTGTCCAAGGAGGAGCTGATCCAGAACATGGAACCGGTGGA 654  
QY 541 CCGAGAGATCAACATGTGTAGAGCAGACAGATCTTAAGCTGAAGAAAGACAGCAACAGCT 600  
DB 655 CCGAGAGATCAACATGTGTAGAGCAGACAGATCTTAAGCTGAAGAAAGACAGCAACAGCT 714  
QY 601 GAGAGAGAGAGGCTGGCCAAAGCCCGGAGCTGTAGAAAGCCCGTGTCAACGCGCCCATGCA 660  
DB 715 GAGAGAGAGAGGCTGGCCAAAGCCCGGAGCTGTAGAAAGCCCGTGTCAACGCGCCCATGCA 774  
QY 661 GTCCGAAGACCGCAGACTGTGTGACATCATCAACAGAGAAACCGGAAGAAAGGCTGAAGC 720  
DB 775 GTCCGAAGACCGCAGACTGTGTGACATCATCAACAGAGAAACCGGAAGAAAGGCTGAAGC 834  
QY 721 TGCAATCGGATTTGTGAAGGCTTGGGGCCCGAGGTGAGCTGCGCTGTACAACAGGC 780  
DB 835 TGCAATCGGATTTGTGAAGGCTTGGGGCCCGAGGTGAGCTGCGCTGTACAACAGGC 894  
QY 781 CTCGCAACATCCGGCAGTATCATGAGAATCAATAAATAAAACAAGCGATGCGGAAGAAAGCT 840  
DB 895 CTCGCAACATCCGGCAGTATCATGAGAATCAATAAATAAAACAAGCGATGCGGAAGAAAGCT 954  
QY 841 AATCTTTGTAATTCAGAAGAGAAATCAAGCTTGGAAACAAATGGAACAGAAATTTCTGCA 900  
DB 955 AATCTTTGTAATTCAGAAGAGAAATCAAGCTTGGAAACAAATGGAACAGAAATTTCTGCA 1014

QY 901 GCGCTATGACCAAGCTCATGAGAGCCTTGGAAAAAAGGTGAGGCCATCGAAAAACAACC 960  
DB 1015 GCGCTATGACCAAGCTCATGAGAGCCTTGGAAAAAAGGTGAGGCCATCGAAAAACAACC 1074  
QY 961 GCGCGCGCGGCGCAAGAGAGAGAGTGGCGGAGTACTACGAAGAAAGCATTCCTCGAGAT 1020  
DB 1075 CCGGCGCGGCGCAAGAGAGAGAGTGGCGGAGTACTACGAAGAAAGCATTCCTCGAGAT 1134  
QY 1021 CCGCAAGCAGCGCGAGCTTGCAGAGAGCGCATGCGAGAGAGGTTGGCCAGCGGGCAGTGG 1080  
DB 1135 CCGCAAGCAGCGCGAGCTTGCAGAGAGCGCATGCGAGAGAGGTTGGCCAGCGGGCAGTGG 1194  
QY 1081 GCTGTTCATGTTCGCGCGCGCCGACGACAGAGGTTCAGAGATCATCATGTGGCTCTC 1140  
DB 1195 GCTGTTCATGTTCGCGCGCGCCGACGACAGAGGTTCAGAGATCATCATGTGGCTCTC 1254  
QY 1141 AGAGCAGAGAAACCTGGAAGACAGATGCGCAGCTGGCCGGATCCGCGCCATGCTGTA 1200  
DB 1255 AGAGCAGAGAAACCTGGAAGACAGATGCGCAGCTGGCCGGATTCGCGCCATGCTGTA 1314  
QY 1201 CGACGCTGACAGCAGCGCATCAAGTTTCATCAACATGAACGGGCTTATGGCCGACCCCAT 1260  
DB 1315 CGACGCTGACAGCAGCGCATCAAGTTTCATCAACATGAACGGGCTTATGGCCGACCCCAT 1374  
QY 1261 GAAGGTGTAACAAGACCGCCAGGTTCATGAACATGTGAGTGAAGAGAAAGAGACCTT 1320  
DB 1375 GAAGGTGTAACAAGACCGCCAGGTTCATGAACATGTGAGTGAAGAGAAAGACCTT 1434  
QY 1321 CCGGGAAGAAATTCAATGACAGATCCCAAGACTTTGGCCGATGCGATCATCTTGGAGAG 1380  
DB 1435 CCGGGAAGAAATTCAATGACAGATCCCAAGACTTTGGCCGATGCGATCATCTTGGAGAG 1494  
QY 1381 GAAGCAGTGTGCTGAGTGCCTCTTATTACTACTGACTGAAGAAAGAAATGAACTATA 1440  
DB 1495 GAAGCAGTGTGCTGAGTGCCTCTTATTACTACTGACTGAAGAAAGAAATGAACTATA 1554  
QY 1441 GAGCTGTGTAAGCAGACCTATTCGGCGCCGCGGCAAGGCCAGCAGCAACAACAGACGA 1500  
DB 1555 GAGCTGTGTAAGCAGACCTATTCGGCGCCGCGGCAAGGCCAGCAGCAACAACAGACGA 1614  
QY 1501 GCAGCAGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
DB 1615 GCAGCAGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1674  
QY 1561 AGATGAGAAAGAGAAAGAAAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620  
DB 1675 AGATGAGAAAGAGAAAGAAAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1734  
QY 1621 CGACAAAGGAAGACTCTTCAAGAGAGAGACAGACGACACTTCAAGGAGAGACAAAGCAG 1680  
DB 1735 CGACAAAGGAAGACTCTTCAAGAGAGAGACAGACGACACTTCAAGGAGAGAGCAGAGCAG 1794  
QY 1681 GAAAGAGGCTGTGGCTTCCAAAGGCGCGCAAACTGCGCAACGCGGAGAAAGCGAAAGG 1740  
DB 1795 GAAAGAGGCTGTGGCTTCCAAAGGCGCGCAAACTGCGCAACGCGGAGAAAGCGAAAGG 1854  
QY 1741 CCGCATCAACCGGCTCAATGTGCTTAATGAGGCCAACAGGAGAGGCTATCACCCCCAGCA 1800  
DB 1855 CCGCATCAACCGGCTCAATGTGCTTAATGAGGCCAACAGGAGAGGCTATCACCCCCAGCA 1914  
QY 1801 GAGCGCCGAGGCTGGCTTCATGAGCTGAATGAGATTCTCGCTGGAAGAAAGAAAT 1860  
DB 1915 GAGCGCCGAGGCTGGCTTCATGAGCTGAATGAGATTCTCGCTGGAAGAAAGAAAT 1974  
QY 1861 GGAACACGCAAGAAAGTCTTCTTGAAACAAGGCGCAACTGTGTGGCCATTCGCGCGAT 1920  
DB 1975 GGAACACGCAAGAAAGTCTTCTTGAAACAAGGCGCAACTGTGTGGCCATTCGCGCGAT 2034  
QY 1921 GGTGGGCTCAAGACTGTGTGCGAGGTGAAGAACTTCACTTCAACTCAAGAAAGGCA 1980  
DB 2035 GGTGGGCTCAAGACTGTGTGCGAGGTGAAGAACTTCACTTCAACTCAAGAAAGGCA 2094  
QY 1981 GAACCTCGATGAGATCTTGCAGCAGCAACGCTGAAGATGGAAGAGAGAGACGCGCG 2040

Db 2095 GAACCTCGATGATCTTGACAGCACAAGCTGAAGATGAGAAAGAGAAACCGCG 2154  
Qy 2041 GAGGAAGAAAGAAAGACCGCGCGCGCGCGCAAGCGAGAGAGCTGCAATCCCGCCCTGGT 2100  
Db 2155 GAGGAAGAAAGAAAGACCGCGCGCGCGCGCAAGCGAGAGAGCTGCAATCCCGCCCTGGT 2214  
Qy 2101 GAGGATGAGAGATGAGAGCGCTCGGCGCTGAGCGGAAATGAGAGAGAGATGGTGGAGA 2160  
Db 2215 GAGGATGAGAGATGAGAGCGCTCGGCGCTGAGCGGAAATGAGAGAGAGATGGTGGAGA 2274  
Qy 2161 GGCTGAAGCCTTACATGCTCTTGGAATGAGGTGCCAAGGGGAAATGCAATGCGCCAGC 2220  
Db 2275 GGCTGAAGCCTTACATGCTCTTGGAATGAGGTGCCAAGGGGAAATGCAATGCGCCAGC 2334  
Qy 2221 CACTGTCAACAACAGTCTGACACCGAGAGCAATCCCTCTCTCTCAACATGAGGCCCGCA 2280  
Db 2335 CACTGTCAACAACAGTCTGACACCGAGAGCAATCCCTCTCTCTCAACATGAGGCCCGCA 2394  
Qy 2281 GAGACAGGGGCAAGATGGCGCCCAAGCGCCCAAGCCTTGCGCGCGCGAGCGGCAACCCG 2340  
Db 2395 GAGACAGGGGCAAGATGGCGCCCAAGCGCCCAAGCCTTGCGCGCGCGAGCGGCAACCCG 2454  
Qy 2341 AGGCCCCACCCCAACAAGAGACATCCGCGGCCCAATTTGAGCCCAACCCCGGCTTC 2400  
Db 2455 AGGCCCCACCCCAACAAGAGACATCCGCGGCCCAATTTGAGCCCAACCCCGGCTTC 2514  
Qy 2401 TGAAGCAACCGGAGCCCTTACGCGCCCAACAGACCCCAACGCTCTGTGCACTCTTCG 2460  
Db 2515 TGAAGCAACCGGAGCCCTTACGCGCCCAACAGACCCCAACGCTCTGTGCACTCTTCG 2574  
Qy 2461 TGTGTGCCCCAAGAGAGAGAGAGAGAGACCGAGAGCGGCCCAAGTGGAGAGAG 2520  
Db 2575 TGTGTGCCCCAAGAGAGAGAGAGAGAGAGACCGAGAGCGGCCCAAGTGGAGAGAG 2634  
Qy 2521 GAGAGAGCAAGAGCCCGCGCGCTGAGAGCTGCGAGTGAACAAGGAAAGCCGAGGA 2580  
Db 2635 GAGAGAGCAAGAGCCCGCGCGCTGAGAGCTGCGAGTGAACAAGGAAAGCCGAGGA 2694  
Qy 2581 GCCCGTCAAGAGAGAGTGAACCGAGAGAACCGAGAGGGGCGCGCAAGGCAAGAGCCG 2640  
Db 2695 GCCCGTCAAGAGAGAGTGAACCGAGAGAACCGAGAGGGGCGCGCAAGGCAAGAGCCG 2754  
Qy 2641 GAGAGCGCTGAGAGCAAGCGCGAGGGGCGCTCAAGGCAAGAGAGAGAGGGCGAGAG 2700  
Db 2755 GAGAGCGCTGAGAGCAAGCGCGAGGGGCGCTCAAGGCAAGAGAGAGAGGGCGAGAG 2814  
Qy 2701 CGGCAAGGCGCACACTGCGCAAGAGCTCGGAGCGGCCCGCAAGACAGGACTCCAGTCTAC 2760  
Db 2815 CGGCAAGGCGCACACTGCGCAAGAGCTCGGAGCGGCCCGCAAGACAGGACTCCAGTCTAC 2874  
Qy 2761 CTGCAAGTGAACAGAGTGAATGAGGCGAGGGCGCGCAAGAACCGGCTGTGTCCC 2820  
Db 2875 CTGCAAGTGAACAGAGTGAATGAGGCGAGGGCGCGCAAGAACCGGCTGTGTCCC 2934  
Qy 2821 AAGGCGCAGGCTCTCAACCGCACTGGGAGACCGCGGGGCAATGCTCAACCCAGAGAGC 2880  
Db 2935 AAGGCGCAGGCTCTCAACCGCACTGGGAGACCGCGGGGCAATGCTCAACCCAGAGAGC 2994  
Qy 2881 ACTGGAAGCTGGAAGAGCTGGAAGCAGGAGCGGCTGCATCCCGCCATCCAGGTCAACAA 2940  
Db 2995 ACTGGAAGCTGGAAGAGCTGGAAGCAGGAGCGGCTGCATCCCGCCATCCAGGTCAACAA 3054  
Qy 2941 AGTCCATGAGCCCCCGGAGAGAGCAGAGCTCCCAAGAGGCTCCCGCAGCCGACCC 3000  
Db 3055 AGTCCATGAGCCCCCGGAGAGAGCAGAGCTCCCAAGAGGCTCCCGCAGCCGACCC 3114  
Qy 3001 GCCACGCAAAACCTGCAAGCGGAGAGAGCGCCCTCAAGAGGCTGGAGAGAGAGCCCGG 3060  
Db 3115 GCCACGCAAAACCTGCAAGCGGAGAGAGCGCCCTCAAGAGGCTGGAGAGAGAGCCCGG 3174  
Qy 3061 GGGCAAGAGAGAGAGCCCGGCAACCCCGCGCAAGAGAGGCTTGCAGCGGAGGCCCA 3120

Db 3175 GGGCAAGAGAGAGAGCCCGGCAACCCCGCGCAAGAGAGGCTTGCAGCGGAGGCCCA 3234  
Qy 3121 GAAGCTGCGGGGAGACCCGCTTGTGAGACTTCGGGCTGGCCCTTCCCGTGGCCCCCG 3180  
Db 3235 GAAGCTGCGGGGAGACCCGCTTGTGAGACTTCGGGCTGGCCCTTCCCGTGGCCCCCG 3294  
Qy 3181 TGAAGTATCAAGGCTCCCGCATGCGCCGAGACCCCTCAGGCTTCTCTACGCTCACG 3240  
Db 3295 TGAAGTATCAAGGCTCCCGCATGCGCCGAGACCCCTCAGGCTTCTCTACGCTCACG 3354  
Qy 3241 TGGTCAACCACTGCGCTTGGGCTCCATGACACTGCGCGGCCGCTCTGCGCGCCAGC 3300  
Db 3355 TGGTCAACCACTGCGCTTGGGCTCCATGACACTGCGCGGCCGCTCTGCGCGCCAGC 3414  
Qy 3301 CACCATCTCAACCGGCGTCCCGCTCATCTGTGCAAGCAACCGAGGCTCTCGAGAG 3360  
Db 3415 CACCATCTCAACCGGCGTCCCGCTCATCTGTGCAAGCAACCGAGGCTCTCGAGAG 3474  
Qy 3361 GCAAATAGGTGCATCTCCCAAGGAATGTGGTCCAGTTCAGTCCGTAATCAGAGCA 3420  
Db 3475 GCAAATAGGTGCATCTCCCAAGGAATGTGGTCCAGTTCAGTCCGTAATCAGAGCA 3534  
Qy 3421 TGCCAAGGCCCGGTGGGCTCTGTCACATGCGGCTGCCCCGCAAGAGAGAGAGAG 3480  
Db 3535 TGCCAAGGCCCGGTGGGCTCTGTCACATGCGGCTGCCCCGCAAGAGAGAGAGAG 3594  
Qy 3481 GCTGGCAACCTTCAAGGAGTGAAGAGAGAGAGAGAGTGTCCCAAGGGGCGAGGCTGGGCG 3540  
Db 3595 GCTGGCAACCTTCAAGGAGTGAAGAGAGAGAGAGAGTGTCCCAAGGGGCGAGGCTGGGCG 3654  
Qy 3541 ACCGAGAGGCTGGGGGTGCCCAAGCGCAGAGAGGCGTCCGTGTGAGAGAGAGAGCTCT 3600  
Db 3655 ACCGAGAGGCTGGGGGTGCCCAAGCGCAGAGAGGCGTCCGTGTGAGAGAGAGAGCTCT 3714  
Qy 3601 GGGCTCAGTTCCGGGCGAAGAGATCAACAAGGCAATTCACAGCAACGCGTCTCGGA 3660  
Db 3715 GGGCTCAGTTCCGGGCGAAGAGATCAACAAGGCAATTCACAGCAACGCGTCTCGGA 3774  
Qy 3661 CAGCGCATCAATACCGGGGCTCCATCAACCGCAAGCGGCAAGGCTGCTGTACAA 3720  
Db 3775 CAGCGCATCAATACCGGGGCTCCATCAACCGCAAGCGGCAAGGCTGCTGTACAA 3834  
Qy 3721 GGGCAACATCAACAGGATATCGGCGAGAGACAGCCGAGTCCGTTGAGACCGCGGCGGA 3780  
Db 3835 GGGCAACATCAACAGGATATCGGCGAGAGACAGCCGAGTCCGTTGAGACCGCGGCGGA 3894  
Qy 3781 GGAACAGCTGCGCAAGGCGCACGTATCAAGAGGCAAGAGGCGACGCTTGTCTTA 3840  
Db 3895 GGAACAGCTGCGCAAGGCGCACGTATCAAGAGGCAAGAGGCGACGCTTGTCTTA 3954  
Qy 3841 TGAAGGTGCAATGTCTGTGAACCAATGCTCCAAAGAGAGCGGCAAGAGAGCTCAAGAGC 3900  
Db 3955 TGAAGGTGCAATGTCTGTGAACCAATGCTCCAAAGAGAGCGGCAAGAGAGCTCAAGAGC 4014  
Qy 3901 CCCCCATGAGAGCGGCGCGCCCAAGCGCACTATGACATGAGAGGCGCGGTGGGCGAG 3960  
Db 4015 CCCCCATGAGAGCGGCGCGCCCAAGCGCACTATGAGAGAGGCGCGGTGGGCGAG 4074  
Qy 3961 AGCCATCTCTTAGCAGCAGATGGAAGGTCTATGAGGCGGTGCATCCGCGGAGAGCA 4020  
Db 4075 AGCCATCTCTTAGCAGCAGATGGAAGGTCTATGAGGCGGTGCATCCGCGGAGAGCA 4134  
Qy 4021 CAGCCCCCAACCTCAAGAGAGACACACATCCGCGGCTCATCAACAAGGAGATCCC 4080  
Db 4135 CAGCCCCCAACCTCAAGAGAGACACACATCCGCGGCTCATCAACAAGGAGATCCC 4194  
Qy 4081 TCGGTCTTACGTGAGAGCAAGAGAGACTTACTGAGTGGAGAGGCAAGCTCTTAAAGG 4140  
Db 4195 TCGGTCTTACGTGAGAGCAAGAGAGACTTACTGAGTGGAGAGGCAAGCTCTTAAAGG 4254  
Qy 4141 GAGAGGCAAGCTTCCGCCCCACCGCCCTCAAGGAGCTTGAAGGCTTCAAGAGCGCA 4200  
Db 4255 GAGAGGCAAGCTTCCGCCCCACCGCCCTCAAGGAGCTTGAAGGCTTCAAGAGCGCA 4314

OY 4201 GGCCTTGAGCCCCCTGAAGCTGAAGCGGCCCCATGAGGGCTGTGTGGCCACGGTGAAGA 4260  
DB 4315 GGCCCTGGGGCCCCCTGAAGCTGAAGCGGCCCATGAGGGGCTGTGTGGCCACGGTGAAGA 4374  
OY 4261 GGGCGGGCGCTCATCTCATGAGAATCCCGCGCGAGAGCTGTGGCACAACGCCGAGCTGGC 4320  
DB 4375 GGGCGGGCGCTCATCTCATGAGAATCCCGCGCGAGAGCTGTGGCACAACGCCGAGCTGGC 4434  
OY 4321 CTGGCCCCCGGGCGGCTCAAGAGGGCTCATCAAGCAGGGCACCCGCTCAAGTACGA 4380  
DB 4435 CTTGGCCCCCGGGCGGCTCAAGAGGGCTCATCAAGCAGGGCACCCGCTCAAGTACGA 4494  
OY 4381 CACCGCGCGTCCACCATCTGTCTCAAAAAAGCAGCGTACGCTCTCATCGGAGGCC 4440  
DB 4495 CACCGCGCGTCCACCATCTGTCTCAAAAAAGCAGCGTACGCTCTCATCGGAGGCC 4554  
OY 4441 CGGCGCGAGCTTCCCAACCCGTGTGACCCGCTGGAATGTATGTATGGCCAGCCCGGCACTGGA 4500  
DB 4555 CGGCGCGAGCTTCCCAACCCGTGTGACCCGCTGGAATGTATGTATGGCCAGCCCGGCACTGGA 4614  
OY 4501 ACGTCTGTCTCAAGAGAGCGCTGAAGAGCGGCGCAGGGACCGCAGAGCTCGGGGG 4560  
DB 4615 ACGTCTGTCTCAAGAGAGCGCTGAAGAGCGGCGCAGGGACCGCAGAGCTCGGGGG 4674  
OY 4561 CTTCATTGTGCGCGGCGCGCCCGGTCAATTGTGCTGTGAGCTGGTAAAGCGCGCAGAGCCC 4620  
DB 4675 CTTCATTGTGCGCGGCGCGCCCGGTCAATTGTGCTGTGAGCTGGGCAAGCGCGCAGAGCCC 4734  
OY 4621 CCGTGCATATAGAGACACAGGGGCAACCTTTGGCCGCGCACCTTCCACAGAGTTCCCGCT 4680  
DB 4735 CTTAATCTATAGAGACACAGGGGCAACCTTTGGCCGCGCACCTTCCACAGAGTTCCCGCT 4794  
OY 4681 GACCATGCGGAGGCCCAACGCGCGCTGTGAGAGGAGGACCTTTGCTCCAGAAAGCATC 4740  
DB 4795 GACCATGCGGAGGCCCAACGCGCGCTGTGAGAGGAGGAGGACCTTTGCTCCAGAAAGCATC 4854  
OY 4741 CCAGGACCGAAAGCTGACGTTCGACGCTTGTGAGATTCGCCAAGTCCCGCAGACGCGT 4800  
DB 4855 CCAGGACCGAAAGCTGACGTTCGACGCTTGTGAGATTCGCCAAGTCCCGCAGACGCGT 4914  
OY 4801 GCCCGAGACCCACCCACACCCCATCTCGCCCTATATAGACACTGTCTTCGGGGGTGATGTG 4860  
DB 4915 GCCCGAGACCCACCCACACCCCATCTCGCCCTATATAGACACTGTCTTCGGGGGTGATGTG 4974  
OY 4861 CGTGAACCTGTATCGCAGGCAATCCCTGAGCTTTCGACCCCACTTCATACCCCGGG 4920  
DB 4975 CGTGAACCTGTATCGCAGGCAATCCCTGAGCTTTCGACCCCACTTCATACCCCGGG 5034  
OY 4921 CATCCCTTGTGACGAGCGGCTGTACTTACCTGTGCCCGCAGCACTGTGCCCAACCCAC 4980  
DB 5035 CATCCCTTGTGACGAGCGGCTGTACTTACCTGTGCCCGCAGCACTGTGCCCAACCCAC 5094  
OY 4981 CTACCCCGACCTGTATCCACCCCTTACTCTACCTCGCGGCTTACCCCGACAGCGGCGCTGGA 5040  
DB 5095 CTACCCCGACCTGTATCCACCCCTTACTCTACCTCGCGGCTTACCCCGACAGCGGCGCTGGA 5154  
OY 5041 GAACCGGAGACCATCATATGATGATCTATCACTACCTTCGACAGAGATGACACAAGACGCG 5100  
DB 5155 GAACCGGAGACCATCATATGATGATCTATCACTACCTTCGACAGAGATGACACAAGACGCG 5214  
OY 5101 CACCGCATGTGCGCAGCGAGCTGATATGTGTAGGGGGCTTCTGCGCCCGGAGTCTCGCT 5160  
DB 5215 CACCGCATGTGCGCAGCGAGCTGATATGTGTAGGGGGCTTCTGCGCCCGGAGTCTCGCT 5274  
OY 5161 GGCATCTCAACTACGTGTGGGGTCCCCGAGGATCATTCGACTGTGCCAAGTGCACACTT 5220  
DB 5275 GGCATCTCAACTACGTGTGGGGTCCCCGAGGATCATTCGACTGTGCCAAGTGCACACTT 5334  
OY 5221 GCGTGTGCTGTGCCCCCGACACAGGACCCCGACGACCGGCATGTGAGACCGGCTTGCTTA 5280  
DB 5335 GCGTGTGCTGTGCCCCCGACACAGGACCCCGACGACCGGCATGTGAGACCGGCTTGCTTA 5394

OY 5281 CTTCCCAACCGCGGCCCAACCCCTTTCAGCAACCGGCCACAGACAGCTTCCCATCTTCCCAAG 5340  
DB 5395 CTTCCCAACCGCGGCCCAACCCCTTTCAGCAACCGGCCACAGACAGCTTCCCATCTTCCCAAG 5454  
OY 5341 AGTCCCAACACATTGAATAAAACCAACCAACGCTCTGTGTCCGAGGGGGAGCGAGACCG 5400  
DB 5455 AGTCCCAACACATTGAATAAAACCAACCAACGCTCTGTGTCCGAGGGGGAGCGAGACCG 5514  
OY 5401 GGATCAGAGCGGAGCCGGAGTCCGGAGCGGGAAAAGTCAATCTTCACTCACTCAACCAACAC 5460  
DB 5515 GGATCAGAGCGGAGCCGGAGTCCGGAGCGGGAAAAGTCAATCTTCACTCACTCAACCAACAC 5574  
OY 5461 GGTGAGACACGCACTCATCTGAGACCTGTGTACAGAGACAGACGCGCAGACGCGGAG 5520  
DB 5575 GGTGAGACACGCACTCATCTGAGACCTGTGTACAGAGACAGACGCGCAGACGCGGAG 5634  
OY 5521 CAGCGCGGGGTGTGGGGCACAAGAGCGCGCCCGCTCCCATCTCCATATGCCACACACA 5580  
DB 5635 CAGCGCGGGGTGTGGGGCACAAGAGCGCGCGCCCGCTCCCATCTCCATATGCCACACACA 5694  
OY 5581 CTGCGCCATCTCCCTCGGACCCAGAGATGCCCTCCAGACAGACCCAGTGTTCACAA 5640  
DB 5695 CTGCGCCATCTCCCTCGGACCCAGAGATGCCCTCCAGACAGACCCAGTGTTCACAA 5754  
OY 5641 CACAGGCAATGAAGGTATCATACCGCTGTGAGAGCCACGCAAGCCACGCTCTGAGGTC 5700  
DB 5755 CACAGGCAATGAAGGTATCATACCGCTGTGAGAGCCACGCAAGCCACGCTCTGAGGTC 5814  
OY 5701 CACCTCCACCTCTCACCGGTTGCGCCAGCTGCCACATTCGCCACCTGCAACCACTGCGC 5760  
DB 5815 CACCTCCACCTCTCACCGGTTGCGCCAGCTGCCACATTCGCCACCTGCAACCACTGCGC 5874  
OY 5761 ACTGGCGGCAACCTTCGATGGGGTCTACCTTACCTCATGAGAGCCGTTGTGTCGCCAA 5820  
DB 5875 ACTGGCGGCAACCTTCGATGGGGTCTACCTTACCTCATGAGAGCCGTTGTGTCGCCAA 5934  
OY 5821 GAGAGCCCCCGGGTGTGCCCCGCGCAGAGCGGCCCGCAGACACACCGGCATGCTTCT 5880  
DB 5935 GAGAGCCCCCGGGTGTGCCCCGCGCAGAGCGGCCCGCAGACACACCGGCATGCTTCT 5994  
OY 5881 GCGCAAGCCCCCAGCGGCTCCGGCTGGAAGCGGCGCTCTCCCGCAGCAAGGGCTCGGA 5940  
DB 5995 GCGCAAGCCCCCAGCGGCTCCGGCTGGAAGCGGCGCTCTCCCGCAGCAAGGGCTCGGA 6054  
OY 5941 GCCCGGCCCCTAGTCTCTGTCTGTGCGCAGCGCACATTCGCCGACACCCCTGCGAA 6000  
DB 6055 GCCCGGCCCCTAGTCTCTGTCTGTGCGCAGCGCACATTCGCCGACACCCCTGCGAA 6114  
OY 6001 GAACCTCGACCTTACCAACGCGCAAGCCGAGCCCGCGGCGCACTGTGCTGCGGCTCGGA 6060  
DB 6115 GAACCTCGACCTTACCAACGCGCAAGCCGAGCCCGCGGCGCACTGTGCTGCGGCTCGGA 6174  
OY 6061 CCGGACCGGGAAAAAGTCAAGATGAACCTTTTTCATTCAGGAATCGGAATCCGTTTC 6120  
DB 6175 CCGGACCGGGAAAAAGTCAAGATGAACCTTTTTCATTCAGGAATCGGAATCCGTTTC 6234  
OY 6121 TCTGGTTTACCAACGCGCAGCACTACAGCCCGGAAGGGGTGAGGCCGTCAACCCCTGTGAG 6180  
DB 6235 TCTGGTTTACCAACGCGCAGCACTACAGCCCGGAAGGGGTGAGGCCGTCAACCCCTGTGAG 6294  
OY 6181 CTCAACCAAGTTTACCAACGAGCAAGAGGGCTTCCCAAGACCTGTGAAGCTGTGACAAAG 6240  
DB 6295 CTCAACCAAGTTTACCAACGAGCAAGAGGGCTTCCCAAGACCTGTGAAGCTGTGACAAAG 6354  
OY 6241 CCACCTGAGAGGGAGAGCTGCGGCGCCCAAGCAACGAGCGCGGTGAAGCTTGTGCGGGGAGGC 6300  
DB 6355 CCACCTGAGAGGGAGAGCTGCGGCGCCCAAGCAACGAGCGCGGTGAAGCTTGTGCGGGGAGGC 6414  
OY 6301 CGCCCACTTCCCAACCTTGTGCGGCGCTGTGAGAGCCAGGACCTGTTCACAGCCGCTGCT 6360  
DB 6415 CGCCCACTTCCCAACCTTGTGCGGCGCTGTGAGAGCCAGGACCTGTTCACAGCCGCTGCT 6474  
OY 6361 CCAAGCGGCCCCAGGGGTCAAAAGTCAACAGCGGGGTGTCACTCTGGCCAGCACATCAG 6420



Db	6475	CGAGACGGCCCAAGGAGTCAAAAGGTCAACACAGGAGTGTCAACCTTGCCCAAGCAATACAG	6534
Qy	6421	TGAGGTCAATCAACAGAGACTACACCCGGCAACACCCAGAGAGTCAAGCGCAACCCCTGCGC	6480
Db	6535	TGAGGTCAATCAACAGAGACTACACCCGGCAACACCCAGAGAGTCAAGCGCAACCCCTGCGC	6594
Qy	6481	CGCCCCCTCTAATCTCTTCCCTGGGGGCGAGTGGCCCGTCTTGGAACCTTCGGCGCCCAAC	6540
Db	6595	CGCCCCCTCTAATCTCTTCCCTGGGGGCGAGTGGCCCGTCTTGGAACCTTCGGCGCCCAAC	6654
Qy	6541	CAGTGAACCTTAACCTCCCGGCCCCGGGACCAATGATGCCCCGGGCGCTGCTCCCGCAAG	6600
Db	6655	CAGTGAACCTTAACCTCCCGGCCCCGGGACCAATGATGCCCCGGGCGCTGCTCCCGCAAG	6714
Qy	6601	CGAAGGGGGGCAAGAGGTCTCCAGAGCCCAACAAAGACGTCCGCTTTGGGTGTGTGAGGA	6660
Db	6715	CGAAGGGGGGCAAGAGGTCTCCAGAGCCCAACAAAGACGTCCGCTTTGGGTGTGTGAGGA	6774
Qy	6661	CGGTAATTGAACCTGTGTCCCGACCGGAGGGGATGAACGAGCCAGAGGCACTCCCGAGTGC	6720
Db	6775	CGGTAATTGAACCTGTGTCCCGACCGGAGGGGATGAACGAGCCAGAGGCACTCCCGAGTGC	6834
Qy	6721	TGTGTATCCCGCTGCTGTACCGGGAATGGGGAAACAACGGAAGCCCAAGCAGAGATGGGCTCCA	6780
Db	6835	TGTGTATCCCGCTGCTGTACCGGGAATGGGGAAACAACGGAAGCCCAAGCAGAGATGGGCTCCA	6894
Qy	6781	GTCTCCAGGGCAACACAGCCAGAGCCGGCCAGCTCTTTCAGCAAGGCTGACCCGAGCAACTC	6840
Db	6895	GTCTCCAGGGCAACACAGCCAGAGCCGGCCAGCTCTTTCAGCAAGGCTGACCCGAGCAACTC	6954
Qy	6841	CGCCATGGTCAAGTGTCCAAAGAACCAAGAGATCAACAAGAGCTGAACAACCAACCGGAA	6900
Db	6955	CGCCATGGTCAAGTGTCCAAAGAACCAAGAGATCAACAAGAGCTGAACAACCAACCGGAA	7014
Qy	6901	TGAGCTGTGAATACAAATATCAAGCCAGCTGGGACGGAAGTCTTCAATATGCCCGCAATCAC	6960
Db	7015	TGAGCTGTGAATACAAATATCAAGCCAGCTGGGACGGAAGTCTTCAATATGCCCGCAATCAC	7074
Qy	6961	CGGAACAGGGCCTTTATGACTTAATGAAGCGAAGCGGTGACGAACAATGCACAGCAACAACAT	7020
Db	7075	CGGAACAGGGCCTTTATGACTTAATGAAGCGAAGCGGTGACGAACAATGCACAGCAACAACAT	7134
Qy	7021	GGGGCTGAGAGCCAAATATTAGAAAAGCACTCATGGGTAAATATATGACAGTGGGAGAGATC	7080
Db	7135	GGGGCTGAGAGCCAAATATTAGAAAAGCACTCATGGGTAAATATATGACAGTGGGAGAGATC	7194
Qy	7081	CCCGCCGCTCAAGCCGCAATGCTTTTAAACCTCTGATATCCAGTGCAGGCTGCGCGCTGC	7140
Db	7195	CCCGCCGCTCAAGCCGCAATGCTTTTAAACCTCTGATATATCCAGTGCAGGCTGCGCGCTGC	7254
Qy	7141	TATGCCCAATAACCGCTGTGTGAACGAGCGGAGTGAACAACAATTCAACCTGCCAGGTGACGG	7200
Db	7255	TATGCCCAATAACCGCTGTGTGAACGAGCGGAGTGAACAACAATTCAACCTGCCAGGTGACGG	7314
Qy	7201	CGGGAAGGCAAGGCTCTGTGCAGAGCCAGGAGCGAAGAACCAAGTCCCGGCGCCGGG	7260
Db	7315	CGGGAAGGCAAGGCTCTGTGCAGAGCCAGGAGCGAAGAACCAAGTCCCGGCGCCGGG	7374
Qy	7261	CTTGCGCATCTGGGGAACCGGCAACCCCTCTGTCTCTCACTGATGCACTCGAGGGGAGCTGCAA	7320
Db	7375	CTTGCGCATCTGGGGAACCGGCAACCCCTCTGTCTCTCACTGATGCACTCGAGGGGAGCTGCAA	7434
Qy	7321	CCGGCCGACCGCGCTCAACCAACCGCGTGTGGGAAGACAGAGCCCTGTCGCCAGGTTTCAAC	7380
Db	7435	CCGGCCGACCGCGCTCAACCAACCGCGTGTGGGAAGACAGAGCCCTGTCGCCAGGTTTCAAC	7494
Qy	7381	GCCATTCCTCCATCAACCCCTGATCATATCGGCTGCAAGCGGGTGTCAATGCTTCCCAAC	7440
Db	7495	GCCATTCCTCCATCAACCCCTGATCATATCGGCTGCAAGCGGGTGTCAATGCTTCCCAAC	7554
Qy	7441	CCCAACCGGGCTCCCCCGGGCAAGCGGGCCCTCTGCTGGCCCAACAACGCTTGAGACGA	7500

Db	7555	CCCAACCGGGCCTCCCGCGGGCAGAGGGGGCCCTCGCTGGCCCCCAACAACGCTGGAGCA	7611
Oy	7501	GGAGCCCAAGCACTGCTCTGTCTGCAAGTACAGACACTCTCCGACACGAGTACTAG	7560
Db	7615	GGAGCCCAAGCACTGCTCTGTCTGCAAGTACAGACACTCTCCGACACGAGTACTAG	7674
Oy	7561	AAACAAGGCGGGGGGGGGCGGGCGGTBUBRSERLUIGTCAAGTCCACAGAGCCACAGAA	7620
Db	7675	AAACAAGGCGGGGGGGGGCGGGCGG-----TGTCAGGTCCAGAGACCCACAGAA	7724
Oy	7621	CGGCGCTTGCAAGACAGGGGGCGGCTGCGCACTCCCCCAACCAAGAGAGGCCCTGAGTC	7680
Db	7725	CGGCGCTTGCAAGACAGAGGGCGGGCTGCGCACTCCCCCAACCAAGAGAGGCCCTGAGTC	7784
Oy	7681	CGCGTGGCGCTTCATCCATCTGTCCGTCCAGAGCGGCACTCTTGGCTGTCTAAAGCCTT	7740
Db	7785	CGCGTGGCGCTTCATCCATCTGTCCGTCCAGAGCGGCACTCTTGGCTGTCTAAAGCCTT	7844
Oy	7741	AACCTAAGACTCCGCGCCCGGGCTGTGGCTGTGCAGACTTACTAGGGGAGTTTAACTTG	7800
Db	7845	AACCTAAGACTCCGCGCCCGGGCTGTGGCTGTGCAGACTTACTAGGGGAGTTTAACTTG	7904
Oy	7801	GTGCTCCGGGAAAGGAGAGGGGAAAGGGGGCGGGGAGGGGGCAACGACAGGGGTGGACGAC	7860
Db	7905	GTGCTCCGGGAAAGGAGAGGGGAAAGGGGGCGGGGAGGGGGCAACGACAGGGGTGGACGAC	7964
Oy	7861	ACACAGGCGGACAGAGGCGGACAGAGGACCCAAAGCAGAGTGAACAACGACTCCACGCGCAC	7920
Db	7965	ACGCAAGGCGGACAGAGGCGGACAGAGGACCCAAAGCAGAGTGAACAACGACTCCACGCGCAC	8024
Oy	7921	TGCGTCCCCCGAATGCAATTTGGAAACAAAGTCTTAACTGAGCTGCAAGCCCCGCGCCT	7980
Db	8025	TGCGTCCCCCGAATGCAATTTGGAAACAAAGTCTTAACTGAGCTGCAAGCCCCGCGCCT	8084
Oy	7981	CCCGTCCGCGCTCCCATCCCGGCTTAAGCGCTTGGAGACAGTNGACGAGGCGCTGTCAAGCC	8040
Db	8085	CCCGTCCGCGCTCCCATCCCGGCTTAAAGCGCTTGGAGACAGTNGACGAGGCGCTGTCAAGCC	8144
Oy	8041	CCAGTGGCGCTGTTCGGGATCCCAACAAGCTGCCCAAGCAGATTTGCTGGAAACAA	8100
Db	8145	CCAATGGCGCTGTTCGGGATCCCAACAAGCTGCCCAAGCAGATTTGCTGGAAACAA	8204
Oy	8101	GTCAGGCGAGTGGGGCGAACAAAGGGCCAGGTGCGGCTTGGGGGGAAACGAGTCTCCGA	8160
Db	8205	GTCAGGCGAGTGGGGCGAACAAAGGGCCAGGTGCGGCTTGGGGGGAAACGAGTCTCCGA	8264
Oy	8161	GGACTGGAATGTTTTTTTTCACACTCTGTTCGGGAGGGTGGGAAAGGAAGGCGAGTGA	8220
Db	8265	GGACTGGAATGTTTTTTTTCACACTCTGTTCGGGAGGGTGGGAAAGGAAGGCGAGTGA	8324
Oy	8221	AATGATGTTGTGGTTTACAGGGTATTATTTTGTGATACCTTCAATGAATTAATTGAGTGT	8280
Db	8325	AATGATGTTGTGGTTTACAGGGTATTATTTTGTGATACCTTCAATGAATTAATTGAGTGT	8384
Oy	8281	TTACGCAAGAGGACTTACCCAGATATTACTGTGCTGTGTGCTTTTGAATCTCTGTGCTTACG	8340
Db	8385	TTACGCAAGAGGACTTACCCAGATATTACTGTGCTGTGTGCTTTTGAATCTCTGTGCTTACG	8444
Oy	8341	TTCAAGAGGCGGTGTGCAAGCGCCGACAGTCCGATACCTTCCGACAGACCAAGGGGGC	8400
Db	8445	TTCAAGAGGCGGTGTGCAAGCGCCGACAGTCCGATACCTTCCGACAGACCAAGGGGGC	8504
Oy	8401	GGGGAATGCTCG--TCAAGCCCGCGTGTGTCTCTCCCTCCCTTCCCTTGGGGCGAATGA	8459
Db	8505	GGGGAATGCTGACTCAAGCCCGCGTGTGTCTCTCCCTCCCTTCCCTTGGGGCGAATGA	8564
Oy	8460	ATTGATGTCGTAATCTGTGGCCGCAATTTGCGACAGGTTGGTGTGTAATCTGTGATTTACAC	8519
Db	8565	ATTGATGTCGTAATCTGTGGCCGCAATTTGCGACAGGTTGGTGTGTAATCTGTGATTTACAC	8624
Oy	8520	ACGTCGTTCTTAATTAAGAGATTTTACTCCAAAAAAA 8561	
Db	8625	ACGTCGTTCTTAATTAAGAGATTTTACTCCAGTTACAA 8666	

RESULT 4  
US-09-819-104A-1  
; Sequence 1, Application US/09819104A  
; Publication No. US20030027137A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, J. Don  
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES  
; FILE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: UMG-030  
; CURRENT APPLICATION NUMBER: US/09/819,104A  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: 60/193,138  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 8686  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (157)..(7677)  
; US-09-819-104A-1

ABM

Query Match 97.2%; Score 8324.6; DB 3; Length 8686;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 8472; Conservative 1; Mismatches 34; Indels 88; Gaps 4;

QY 1 CATGTGGGCTTCACACAGCTTTGTG3CACAGCTGAGAGGCGCACTGAGCCCGCTACCC 60  
DB 156 CATGTGGGCTTCACACAGCTTTGTG3CACAGCTGAGAGGCGCACTGAGCCCGCTACCC 215  
QY 61 GCGCCACAGCCCTTCTCTCCCGAGTGCAGATGCGCCG3ACGCAACAGGACGTCGCGCTCT 120  
DB 216 GCGCCACAGCCCTTCTCTCCCGAGTGCAGATGCGCCG3ACGCAACAGGACGTCGCGCTCT 275  
QY 121 GAGATGACAGACACACTCCCGGACTATGCTCCCACTGTCGCGGCTCCATCATCA 180  
DB 276 GAGATGACAGACACACTCCCGGACTATGCTCCCACTGTCGCGGCTCCATCATCA 335  
QY 181 GCGCCAGCGGCGAGAGCCCTCTCTGCTGTGAGTTCCAGCCCGGAAATGAACGGTCCA 240  
DB 336 GCGCCAGCGGCGAGAGCCCTCTCTGCTGTGAGTTCCAGCCCGGAAATGAACGGTCCA 395  
QY 241 GAGGCTCCACCTGCGGCGAGAGTCCACATCATCTGCGCGAGCTGGGGAATGAGAT 300  
DB 396 GAGGCTCCACCTGCGGCGAGAGTCCACATCATCTGCGCGAGCTGGGGAATGAGAT 455  
QY 301 GAGGTTCAATTGAAACAGAGCGCTCTGAGTGAAGCTGCTGACCCCTGCTGAGAC 360  
DB 456 GAGGTTCAATTGAAACAGAGCGCTCTGAGTGAAGCTGCTGACCCCTGCTGAGAC 515  
QY 361 GTACCCCTGCTGCGCACGCGGCTCTGCGGAACTGAAGACTTACCAAGAGACCGTAG 420  
DB 516 GTACCCCTGCTGCGCACGCGGCTCTGCGGAACTGAAGACTTACCAAGAGACCGTAG 575  
QY 421 CCTGACGCGGCAAGCTGGAACCGGTGTCTCCCGGAGCCCGCGACACTGACCCCTGAACT 480  
DB 576 CCTGACGCGGCAAGCTGGAACCGGTGTCTCCCGGAGCCCGCGACACTGACCCCTGAACT 635  
QY 481 GAGGCTGTGCGCGCACGCGGCTCTGCAAGAGAGCTGATCCAGAACATG3ACCGGTGGA 540  
DB 636 GAGGCTGTGCGCGCACGCGGCTCTGCAAGAGAGCTGATCCAGAACATG3ACCGGTGGA 695  
QY 541 CCGAGAGATCAACCATGTGAGAGACAGACTCTTAAGCTGAAGAAAGACAGCAACACT 600  
DB 696 CCGAGAGATCAACCATGTGAGAGACAGACTCTTAAGCTGAAGAAAGACAGCAACACT 755  
QY 601 GAGAGAGAGAGCTGCGCAAGCCCGCGAGCTGGAAGAGCCGCTGTACACCGCGCCATGGA 660  
DB 756 GAGAGAGAGAGCTGCGCAAGCCCGCGAGCTGGAAGAGCCGCTGTACACCGCGCCATGGA 815

QY 661 GTGGAAGCACCGCAGCCTGTGTGACATCATACGACGAGAACCGGAAAGAGCTGAAGC 720  
DB 816 GTGGAAGCACCGCAGCCTGTGTGACATCATACGACGAGAACCGGAAAGAGCTGAAGC 875  
QY 721 TGCACATGGAATTTGGAAGGCTTGGGGCCCCAGGTGAGCTGCCGCTGTACACACACC 780  
DB 876 TGCACATGGAATTTGGAAGGCTTGGGGCCCCAGGTGAGCTGCCGCTGTACACACACC 935  
QY 781 CTCGACACCCCGGAGATCATGAGAAATCAATCAAAATAACAGGGGATGCGGAAGACT 840  
DB 936 CTCGACACCCCGGAGATCATGAGAAATCAATCAAAATAACAGGGGATGCGGAAGACT 995  
QY 841 AATCTTGACTTCAAGAGAGAAATCAAGCTTCGGAACATGGAAGCAGAAATTCTGCA 900  
DB 996 AATCTTGACTTCAAGAGAGAAATCAAGCTTCGGAACATGGAAGCAGAAATTCTGCA 1055  
QY 901 GCGCTATGACCAAGCTCATGAGAGCCTTGGAAAAAAGGTGAGCGCATCGAAAAACACC 960  
DB 1056 GCGCTATGACCAAGCTCATGAGAGCCTTGGAAAAAAGGTGAGCGCATCGAAAAACACC 1115  
QY 961 GCGCGCGCGGCGCAAGAGAGCAAGTGCAGAGTACTACGAAAAAGCAGTTCCCTGAGAT 1020  
DB 1116 GCGCGCGCGGCGCAAGAGAGCAAGTGCAGAGTACTACGAAAAAGCAGTTCCCTGAGAT 1175  
QY 1021 CCGCAAGCAGCGCGAGCTTCAGAGACGCGCATGACAGAGGTTGGGCCAGCGGGCAGTGG 1080  
DB 1176 CCGCAAGCAGCGCGAGCTTCAGAGACGCGCATGACAGAGGTTGGGCCAGCGGGCAGTGG 1232  
QY 1081 GCTGTCCATGTTCGCGCGCGCGCGAGACAGAGGTTCAGAGATCATGATGGCTCTTC 1140  
DB 1233 GCTGTCCATGTTCGCGCGCGCGCGAGACAGAGGTTCAGAGATCATGATGGCTCTTC 1292  
QY 1141 AGAGCAGGAGAACTGAGAGAGCAGATGCGCAGCTGSCGTGATCCGCGCATCTGTA 1200  
DB 1293 AGAGCAGGAGAACTGAGAGAGCAGATGCGCAGCTGSCGTGATCCGCGCATCTGTA 1352  
QY 1201 CGAGCTGACCGACGCGCATCAAGTTTCATCAACATGAACGGGCTTATGGCCGACCCAT 1260  
DB 1353 CGAGCTGACCGACGCGCATCAAGTTTCATCAACATGAACGGGCTTATGGCCGACCCAT 1412  
QY 1261 GAAAGGTGACAAAGACCCCGGCTCATGAACTGAGAGTGAAGAGAAAGAGACCTT 1320  
DB 1413 GAAAGGTGACAAAGACCCCGGCTCATGAACTGAGAGTGAAGAGAAAGAGACCTT 1472  
QY 1321 CCGGAGAAAGTTCAATGACAGATCCCAAGACTTTGGCTGATCGCATATTCCTGGAAG 1380  
DB 1473 CCGGAGAAAGTTCAATGACAGATCCCAAGACTTTGGCTGATCGCATATTCCTGGAAG 1532  
QY 1381 GAAAGCACTGCTGAGTGCCTCTTAATTACTACTGTAATAAGAAATGAATCTATTA 1440  
DB 1533 GAAAGCACTGCTGAGTGCCTCTTAATTACTACTGTAATAAGAAATGAATCTATTA 1592  
QY 1441 GAGCTGTGAGAGGAGACTTTCGCGCGCGGCGACAGACGACGAAACAAACAGACCA 1500  
DB 1593 GAGCTGTGAGAGGAGACTTTCGCGCGCGGCGACAGACGACGAAACAAACAGACCA 1652  
QY 1501 GAGCAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
DB 1653 GAGCAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1712  
QY 1561 AGATGAG 1620  
DB 1713 AGATGAG 1772  
QY 1621 CGACAGAGAGAGCTCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680  
DB 1773 CGACAGAGAGAGCTCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1832  
QY 1681 GAAAGAGCTGTGGCTCAAGAGCCGCAAACTGCCAAAGCAGCGGAGAGAGCGCAAGG 1740  
DB 1833 GAAAGAGCTGTGGCTCAAGAGCCGCAAACTGCCAAAGCAGCGGAGAGAGCGCAAGG 1892

QY 1741 CCGCATCACCCGCTCAATGGCTAATGAGGCCAACAGGAGAGGCCATCACCCCCAGCA 1800  
DB CCGCATCACCCGCTCAATGGCTAATGAGGCCAACAGGAGAGGCCATCACCCCCAGCA 1952  
QY 1801 GAGCGCCGAGCTGGCTCTCAATGGAGCTGAATGAGAGTTTCTCGCTGGAACAAGAAAGAAAT 1860  
DB GAGCGCCGAGCTGGCTCTCAATGGAGCTGAATGAGAGTTTCTCGCTGGAACAAGAAAGAAAT 2012  
QY 1861 GGAACAGCCAGAAAGAGTCTCCCTGGAACAAGCGCCGAACTGGTGGGCAATCGCCGGAT 1920  
DB GGAACAGCCAGAAAGAGTCTCCCTGGAACAAGCGCCGAACTGGTGGGCAATCGCCGGAT 2072  
QY 1921 GGTGGGCTCCAGACTGTGTGCGAGTGTAAAGAACTTTCACTTCACTTAAGAAAGAGCA 1980  
DB GGTGGGCTCCAGACTGTGTGCGAGTGTAAAGAACTTTCACTTCACTTAAGAAAGAGCA 2132  
QY 1981 GAACCTCGATGAGATCTTGGACAGCAACAAGCTGAGAGATGAGAGAGAGAGAGAGCGCG 2040  
DB GAACCTCGATGAGATCTTGGACAGCAACAAGCTGAGAGATGAGAGAGAGAGAGAGCGCG 2192  
QY 2041 GAGGAAGAAAGAAAGAGCGCGCGCGCGCGCGCAGCGAGAGAGAGTCCGCGCCGCTG 2100  
DB GAGGAAGAAAGAAAGAGCGCGCGCGCGCGCGCAGCGAGAGAGTCCGCGCCGCTG 2252  
QY 2101 GAGAGATGAGAGATGAGAGCGCTCGGCGCTGAGCGGAAATGAGAGAGAGATGAGTGAAGA 2160  
DB GAGAGATGAGAGATGAGAGCGCTCGGCGCTGAGCGGAAATGAGAGAGAGATGAGTGAAGA 2312  
QY 2161 GGTGAAGCTTCAATGCTCTGGAGATGAGGTGCCAGAGGGAGATGAGTGGCCAGC 2220  
DB GGTGAAGCTTCAATGCTCTGGAGATGAGGTGCCAGAGGGAGATGAGTGGCCAGC 2320  
QY 2221 CACTGTCAACAAGCTCAGACACCGAGAGCATCCCTCTCTCAACTGAGAGCGCGCA 2280  
DB CACTGTCAACAAGCTCAGACACCGAGAGCATCCCTCTCTCAACTGAGAGCGCGCA 2381  
QY 2281 GGAACAAGGACAGATGAGGCCAAGGCCCGCAGCCCTGAGCGCGCAGCGGCAACCC 2340  
DB GGAACAAGGACAGATGAGGCCAAGGCCCGCAGCCCTGAGCGCGCAGCGGCAACCC 2441  
QY 2341 AGGCCCAACCCCAACCGAGAGACATCCGAGGCCCATTTAGGCCCAACCCCGGCTC 2400  
DB AGGCCCAACCCCAACCGAGAGACATCCGAGGCCCATTTAGGCCCAACCCCGGCTC 2442  
QY 2401 TGAAGCACCGGAGCCCTACGCGCCCAACAGACACCCCAATGCCCTCTGCACTCTCC 2460  
DB TGAAGCACCGGAGCCCTACGCGCCCAACAGACACCCCAATTTCTCTTCACTCTCTCC 2502  
QY 2461 TGTGTCTCCCAAG 2520  
DB TGTGTCTCCCAAG 2562  
QY 2521 GAGAGAGCAGAAAGCCCGCGGCTGAGAGCTTGGAGTGAACAAGAGAGAGAGAGAG 2580  
DB GAGAGAGCAGAAAGCCCGCGGCTGAGAGCTTGGAGTGAACAAGAGAGAGAGAGAG 2622  
QY 2581 GCCCGTCAAGAGAGATGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
DB GCCCGTCAAGAGAGATGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2682  
QY 2641 GAGAGCCGCTGAG 2700  
DB GAGAGCCGCTGAG 2742  
QY 2701 CCGCAGAGGACCACTGTCAGAGAGCTCGAGCGCCCCCAAGACAGGAGCTCAAGTCTAC 2760  
DB CCGCAGAGGACCACTGTCAGAGAGCTCGAGCGCCCCCAAGACAGGAGCTCAAGTCTAC 2802  
QY 2761 CTGCAATGACAG 2820  
DB CTGCAATGACAG 2862  
QY 2821 AAGGCCCAAGCTCTCAACCCCGAGCTGGCGAGCCCGCGGCGCAATGCTCAACCCCAAGAGCC 2880

DB 2922 AAGGCCCAAGCTCTCAACCCCGAGCTGGCGAGCCCGCGGCGCAATGCTCAACCCCAAGAGCC 2981  
QY 2881 ACTGACCTGACAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
DB ACTGACCTGACAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2982  
QY 2941 AGTCAATGAGCCCGCGGAG 3000  
DB AGTCAATGAGCCCGCGGAG 3042  
QY 3001 GCCACGCAAAACCTGACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
DB GCCACGCAAAACCTGACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3102  
QY 3061 GGGCAAG 3100  
DB GGGCAAG 3162  
QY 3101 ----GCTTCGACAGCGAG 3156  
DB ----GCTTCGACAGCGAG 3222  
QY 3157 CTTGACCTTCCCGGAG 3216  
DB CTTGACCTTCCCGGAG 3282  
QY 3217 CTGACGCTTCTCTTAAGCTTCAAGTGTCAACGAGAGAGAGAGAGAGAGAGAGAGAGAG 3276  
DB CTGACGCTTCTCTTAAGCTTCAAGTGTCAACGAGAGAGAGAGAGAGAGAGAGAGAGAG 3342  
QY 3277 CCGGCGGCTCTCGAG 3336  
DB CCGGCGGCTCTCGAG 3402  
QY 3337 CAAGCACCCCAAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3396  
DB CAAGCACCCCAAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3462  
QY 3397 GCTTCAGCTCCGTAATCAAGAGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3456  
DB GCTTCAGCTCCGTAATCAAGAGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3522  
QY 3457 GCCCTGCGCAATGAG 3516  
DB GCCCTGCGCAATGAG 3582  
QY 3517 GTCCCAACGAGAGCAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3576  
DB GTCCCAACGAGAGCAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3642  
QY 3577 GTCCGCTGAG 3636  
DB GTCCGCTGAG 3702  
QY 3637 TCCCAAGACAG 3696  
DB TCCCAAGACAG 3762  
QY 3697 CACGCAAGCTGAG 3756  
DB CACGCAAGCTGAG 3822  
QY 3757 GAGTGGCTTGAAGCGGAG 3816  
DB GAGTGGCTTGAAGCGGAG 3882  
QY 3817 CAAG 3876  
DB CAAG 3942  
QY 3942 CAG 4001  
DB CAG 3877  
QY 3877 GAGCGCAG 3936

Db	4002	GGACGGCAGAAAGCAGCTTAGAGACCCCCCATAGAAAGCGGCCGCCCCAAAGCGCACTTAGA	4061
OY	3937	CATGATGAGGGGCGCGCTGTGGCAGAGCCATCTCTCCAGCCAGCATCGAAGCTTCATGGG	3996
Db	4062	CATGATGAGGGGCGCGTGTGGCAGAGCCATCTCTCCAGCCAGCATCGAAGGCTCATGGG	4121
OY	3997	CCGTGGCCATTCGCCGCCGAGGCGACAGAGCCCCCACTCACTGAAAGAGACGCAATCCG	4056
Db	4122	CCGTGCCCAATCCGCCGAGCGACACAGCCCCCACTCAAGAGAGCGACCAATCCG	4181
OY	4057	CGGGTCATCAACAAGAGGATCCCTCGGTCTTAAGTGGAGCAGAGAGACTACCTGCG	4116
Db	4182	CGGGTCATCAACAAGAGGATTCCTCGGTCTTAAGTGGAGCAGAGAGACTACCTGCG	4241
OY	4117	TCGGAGGCCAAGCTCTTAAAGCGGAGGAGCAGCGCTTCGCGCCCAACCGCCCTCAACGAG	4176
Db	4242	TCGGAGGCCAAGCTCTTAAAGCGGAGGAGCAGCGCTTCGCGCCCAACCGCCCTCAACGAG	4301
OY	4177	CCTGACCCGAGGCTCTACAGACGCGAGGCTCTGGGCCCCCTGAAAGCTGAAAGCCGGCCATGA	4236
Db	4302	CCTGACCCGAGGCTCTACAGAGCAGAGGCTCTGGGCCCCCTGAAAGCTGAAAGCCGGCCATGA	4361
OY	4237	GGGCGCTGTGGCCACGCGTGAAGAGAGGCGGGCGCTCCATCCATAGAGATCCCGCGAGAGA	4296
Db	4362	GGGCGCTGTGGCCACGCGTGAAGAGAGGCGGGCGCTCCATCCATAGAGATCCCGCGAGAGA	4421
OY	4297	GCTGCGGACACGCGCCGAGCTGCCCCCTGGGCCCGCGCGCGCTCAAGAGAGGCTCATCAC	4356
Db	4422	GCTGCGGACACGCGCCGAGCTGCCCCCTGGGCCCGCGCGCGCTCAAGAGAGGCTCATCAC	4481
OY	4357	GCAAGGACACCCCGCTCAAGTACGACACCGGCGGCTCCACACTGGCTTCMAAAGACGCA	4416
Db	4482	GCAAGGACACCCCGCTCAAGTACGACACCGGCGGCTCCACACTGGCTTCMAAAGACGCA	4541
OY	4417	CGTACGCTCCCTCAATCGGACAGCCCCCGGCGGAGCGTTCCACCCGCTGACCCCGCTGATGT	4476
Db	4542	CGTACGCTCCCTCAATCGGACAGCCCCCGGCGGAGCGTTCCACCCGCTGACCCCGCTGATGT	4601
OY	4477	GATGCGCAGACGCCCGGAGCACTTGAAAGTGCCTGCTACGAGGAGACCTGAAAGACCGAGC	4536
Db	4602	GATGCGCAGACGCCCGGAGCACTTGAAAGTGCCTGCTACGAGGAGACCTGAAAGACCGAGC	4661
OY	4537	AGGGAACCGCAGCAGACTGGGGGGGCTCCATTGCGGCGGCGCCCCGGTCAATTGTGCTGA	4586
Db	4662	AGGGAACCGCAGCAGACTGGGGGGGCTCCATTGCGGCGGCGCCCCGGTCAATTGTGCTGA	4721
OY	4597	GCTGGGTAAAGCGCGGCGGAGAGCGCCCTGACCTATGAGGACCAAGGGGCAACCTTTGCGG	4656
Db	4722	GCTGGGTAAAGCGCGGCGGAGAGCGCCCTGACCTATGAGGACCAAGGGGCAACCTTTGCGG	4781
OY	4657	CCACCTCCCAAGAGTTGCGCCGTGACCATGCGGAGGCCCAAGCGCGCTGACAGAGGG	4716
Db	4782	CCACCTCCCAAGAGTTGCGCCGTGACCATGCGGAGGCCCAAGCGCGCTGACAGAGGG	4841
OY	4717	CAGCTTTTCGTCAAGCAGGACTCCAGAGACCGAAAGCTGAAGTGAAGCTTCGTGAGAT	4776
Db	4842	CAGCTTTTCGTCAAGCAGGACTCCAGAGACCGAAAGCTGAAGTGAAGCTTCGTGAGAT	4901
OY	4777	CGCCAAGTCCCCCGACAGACCGGTGCGCGAGACCAACCCCAACCCCAATCTGCGCTTGA	4836
Db	4902	CGCCAAGTCCCCCGACAGACCGGTGCGCGAGACCAACCCCAACCCCAATCTGCGCTTGA	4961
OY	4837	GCACCTGCTTGCGGGCGTGAAGTGGAGTGAAGCTGTATGCAAGCACATCCCCCTGGGCTT	4896
Db	4962	GCACCTGCTTGCGGGCGTGAAGTGGAGTGAAGCTGTATGCAAGCACATCCCCCTGGGCTT	5021
OY	4957	CGGAACCTTGCGGCCCAACCCCACTTACCGGACCTGTAAACCAACCCCACTTACCTGCGG	5016
Db	5082	CGGAACCTTGCGGCCCAACCCCACTTACCGGACCTGTAAACCAACCCCACTTACCTGCGG	5141

OY	5017	CTACCCCCGACACGGGCGGCGCTTGGAAGAACGGGAGACCAATCATCAATGACTCATCACTC	5076
Dp	5142	CTACCCCCGACACGGGCGGCGCTTGGAAGAACGGGAGACCAATCATCAATGACTCATCACTC	5201
OY	5077	GCAAGCAGATGACCAACAACAGGGCCACGCGCATGCGCCAGCGAGCTGATGTAGTAGGGGG	5136
Dp	5202	GCAAGCAGATGACCAACAACAGGGCCACGCGCATGCGCCAGCGAGCTGATGTAGTAGGGGG	5261
OY	5137	CTCTCTGCCCCGCGAGTCTCTCGCTGGCACTCACTAATGACGTGGGGTCCCCAGGCATCAT	5196
Dp	5262	CTCTCTGCCCCGCGAGTCTCTCGCTGGCACTCACTAATGACGTGGGGTCCCCAGGCATCAT	5321
OY	5197	CGACCTGTCCCAAGGTGACCACTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5286
Dp	5322	CGACCTGTCCCAAGGTGACCACTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5381
OY	5257	CACCGGCATGACCGGCGCTTGCTTACTCTCCCAACGCGCGCCAGCGCCCTTCAGACGCGCA	5316
Dp	5382	CACCGGCATGACCGGCGCTTGCTTACTCTCCCAACGCGCGCCAGCGCCCTTCAGACGCGCA	5441
OY	5317	CAGGAGCTCCCACTCTCTCCCAAGAGGTTCACAACAATTGACAAAAACAACACACAGTCC	5376
Dp	5442	CAGCAGCTCCCACTCTCTCCCAAGAGGTTCACAACAATTGACAAAAACAACACACAGTCC	5501
OY	5377	CTGTGTCGAGCGGGAGACGAGACCGGGATGAGAGCGGGACCGGGATTCGGAGACGGGAAAA	5436
Dp	5502	CTGTGTCGAGCGGGAGACGAGACCGGGATGAGAGCGGGACCGGGATTCGGAGACGGGAAAA	5561
OY	5437	GTCATCTCTCACGTCACACACGACGAGCGGTGAGGACGCGCACCTATCTGTGAGACTGTGTACGA	5496
Dp	5562	GTCATCTCTCACGTCACACACGAGGTGAGGACGCGCACCTATCTGTGAGACTGTGTACGA	5621
OY	5497	GCGAGAGACGGGTGAGAGAGGGGACGAGCGCGGGGGGTGGGGGACAGACACCGCCCGCG	5556
Dp	5622	GCGAGAGAGGGGTGAGAGAGGGGACGAGCGCGGGGGGTGGGGGACAGACACCGCCCGCG	5681
OY	5557	CTGCCACTCCCAATGAGCCACACGACCTGCGGCCATCTCCCTCGAGCCACAGATGCGCTTCA	5616
Dp	5682	CTGCCACTCCCAATGAGCCACACGACCTGCGGCCATCTCCCTCGAGCCACAGATGCGCTTCA	5741
OY	5617	GCAAGAGCCACAGTGTGCTTCAACAACAAGGACATGAAGGGTATCATCACCGCTGTGGAGCC	5676
Dp	5742	GCAAGAGCCACAGTGTGCTTCAACAACAAGGACATGAAGGGTATCATCACCGCTGTGGAGCC	5801
OY	5677	CAGCAGGCCACGSGTCTTGAAGTTCACCTCCACCTCTCCACCCGTTGCGCCAGCGTGCAC	5736
Dp	5802	CAGCAGGCCACGSGTCTTGAAGTTCACCTCCACCTCTCCACCCGTTGCGCCAGCGTGCAC	5861
OY	5737	ATTGCCACCTGACCACTGCGCCACTGGGCGGACCCCTCATGTGGAGTCTACCTTACCTT	5796
Dp	5862	ATTGCCACCTGACCACTGCGCCACTGGGCGGACCCCTCATGTGGAGTCTACCTTACCTT	5921
OY	5797	CATGAGAGCCGCTTGTCTGCTCCACAAGAGGCGCCCGGGGTGCGCCGGGACAGAGCGGCGCG	5856
Dp	5922	CATGAGAGCCGCTTGTCTGCTCCACAAGAGGCGCCCGGGGTGCGCCGGGACAGAGCGGCGCG	5981
OY	5857	AGCAGACACCGGCGCAAGCCCTTCCTGCGCAAGCGCCCAACCGCGCTCCCGGGCTGGAGCCCGC	5916
Dp	5982	AGCAGACACCGGCGCAAGCCCTTCCTGCGCAAGCGCCCAACCGCGCTCCCGGGCTGGAGCCCGC	6041
OY	5917	CTCCTCCCCCAGCAGAGGCTCGGAGCGCCCGGCGCCCTAGTGTCTCTGTCTGTGGCAAGC	5976
Dp	6042	CTCCTCCCCCAGCAGAGGCTCGGAGCGCCCGGCGCCCTAGTGTCTCTGTCTGTGGCAAGC	6101
OY	5977	CACCATGCGCGCAGACCCCTTGCGAAGAACTCGCACTCAACAACGCAACCGCGAACCCGCGC	6036
Dp	6102	CACCATGCGCGCAGACCCCTTGCGAAGAACTCGCACTCAACAACGCAACCGCGAACCCGCGC	6161
OY	6037	GCGGCAACCTGCTGGGCGCTCGGACCCCGGACCGGGAAAAAAGTCAAAAGTAAACCTTTTC	6096
Dp	6162	GCGGCAACCTGCTGGGCGCTCGGACCCCGGACCGGGAAAAAAGTCAAAAGTAAACCTTTTC	6221

QY 6097 CATCCAGAACTGGAACTCTGTTCTCTGGGTTACACGGGACAGCTTACAGCCCCGGAAG 6156  
DB 6222 CATCCAGAACTGGAACTCTGTTCTCTGGGTTACACGGGACAGCTTACAGCCCCGGAAG 6281  
QY 6157 GGTGGAGCCCGTCAAGCCCTGTGAGCTCAACCAGTGTGACCCAGACAAAGAGGCTCCCA 6216  
DB 6282 GGTGGAGCCCGTCAAGCCCTGTGAGCTCAACCAGTGTGACCCAGACAAAGAGGCTCCCA 6341  
QY 6217 GCACTTGAAGAGCTGACAAAGAGCACTTGAAGGGAGCTGGCCCAAGAGCAGG 6276  
DB 6342 GCACTTGAAGAGCTGACAAAGAGCACTTGAAGGGAGCTGGCCCAAGAGCAGG 6401  
QY 6277 CCCCGTGAAGCTTGGGGGAGGCGCCCACTTCCACACTTGGGCGCTGCTTGAAG 6336  
DB 6402 CCCCGTGAAGCTTGGGGGAGGCGCCCACTTCCACACTTGGGCGCTGCTTGAAG 6461  
QY 6337 CCAAGCCCTGTCCAGCCCGCTGTCTCAAGACCGGCCCAAGGGGTCAAGGTCAAGCGGGT 6396  
DB 6462 CCAAGCCCTGTCCAGCCCGCTGTCTCAAGACCGGCCCAAGGGGTCAAGGTCAAGCGGGT 6521  
QY 6397 GGTCACTCTGSCCAGACACTAGTAGGTCACTACACAGGACTACACCCGACCAACC 6456  
DB 6522 GGTCACTCTGSCCAGACACTAGTAGGTCACTACACAGGACTACACCCGACCAACC 6581  
QY 6457 ACAGCAGCTCAAGGCAACCCCTGCGCCGCCCTCTACTCTCTTCCCTGGGGCAGCTGCC 6516  
DB 6582 ACAGCAGCTCAAGGCAACCCCTGCGCCGCCCTCTACTCTCTTCCCTGGGGCAGCTGCC 6641  
QY 6517 CGTCTTGAACCTTCCGCGCCCAACCAAGTACCTTCACTTCCGCGCCCGGACCAATGCTGC 6576  
DB 6642 CGTCTTGAACCTTCCGCGCCCAACCAAGTACCTTCACTTCCGCGCCCGGACCAATGCTGC 6701  
QY 6577 CCGGCGCCGCTGCTCCCCCAAGCGAAGGGGAGGAGGTCTCCAGGCAACAAAGAC 6636  
DB 6702 CCGGCGCCGCTGCTCCCCCAAGCGAAGGGGAGGAGGTCTCCAGGCAACAAAGAC 6761  
QY 6637 GTTCGCTCTTGGGTGTGTGAGGACGGTATTGAACCTGTGTCCCAACCGGAGGGCAATGAC 6696  
DB 6762 GTTCGCTCTTGGGTGTGTGAGGACGGTATTGAACCTGTGTCCCAACCGGAGGGCAATGAC 6821  
QY 6697 GGAAGCCAGGCACTCCCGAGTGTGTGACCCGCTGTGTACCGGAGTGGGAGCAAGAC 6756  
DB 6822 GGAAGCCAGGCACTCCCGAGTGTGTGACCCGCTGTGTACCGGAGTGGGAGCAAGAC 6881  
QY 6757 GGAAGCCAGGAGTGGGCTCCAGTCTCCAGGCAACCAAGGCAACCGCCAGCCTTCTT 6816  
DB 6882 GGAAGCCAGGAGTGGGCTCCAGTCTCCAGGCAACCAAGGCAACCGCCAGCCTTCTT 6941  
QY 6817 CAGCAAGCTGACCGAGAGCACTCCGCACTGCTCAAGTCCAAAGAACAGAGATCAACA 6876  
DB 6942 CAGCAAGCTGACCGAGAGCACTCCGCACTGCTCAAGTCCAAAGAACAGAGATCAACA 7001  
QY 6877 GAAAGCTGAACCCCAACCCGGAATGAGCTGAATACAATATCAGCCAGCTTGGAGCGGA 6936  
DB 7002 GAAAGCTGAACCCCAACCCGGAATGAGCTGAATACAATATCAGCCAGCTTGGAGCGGA 7061  
QY 6937 GATCTTCAATATGCCCCGCACTACCCGGAACAGGCTTATGACCTTATGAAGCCAGCGGT 6996  
DB 7062 GATCTTCAATATGCCCCGCACTACCCGGAACAGGCTTATGACCTTATGAAGCCAGCGGT 7121  
QY 6997 GAGAGAACATGCGACCAACATGAGGCTGAGAGCATTAATTAAGAAAGCACTCATGGG 7056  
DB 7122 GAGAGAACATGCGACCAACATGAGGCTGAGAGCATTAATTAAGAAAGCACTCATGGG 7181  
QY 7057 TAAATATGACAGTGGGAAGAGTCCCGCGCTCAGCGCCAAATGCTTTTAAACCTCTGA 7116  
DB 7182 TAAATATGACAGTGGGAAGAGTCCCGCGCTCAGCGCCAAATGCTTTTAAACCTCTGA 7241  
QY 7117 TGGCAGTGCAGGCTGCGCTGCTATGCCCCAATAACCGCTGTGAAGGACGAGTGAACA 7176  
DB 7242 TGGCAGTGCAGGCTGCGCTGCTATGCCCCAATAACCGCTGTGAAGGACGAGTGAACA 7301  
QY 7177 CACACTCACTGCGCAGGTGCGGGCGGGAAGGCCAAGGTCTGTGAGAGCCAGCAGCGG 7236

DB 7302 CACACTCACTGCGCAGGTGCGGGCGGGAAGGCCAAGGTCTGTGAGAGCCAGCAGCGG 7361  
QY 7237 AAAAGCAAGTCCCGGACCCCGGGCTGGGACTTGGGGAACCGGCAACCTCTGTCTTC 7296  
DB 7362 AAAAGCAAGTCCCGGACCCCGGGCTGGGACTTGGGGAACCGGCAACCTCTGTCTTC 7421  
QY 7297 AGTGCATCTGGAGGAGAGCTGCAACCGCGGAGCGCGCTCAACCAACCGGCTGTGGAGGA 7356  
DB 7422 AGTGCATCTGGAGGAGAGCTGCAACCGCGGAGCGCGCTCAACCAACCGGCTGTGGAGGA 7481  
QY 7357 CAGGCTCTGTCGCGAGGTTTCCAGCCATTCCTTCAACCCCTGATATCATGTGGCTGCA 7416  
DB 7482 CAGGCTCTGTCGCGAGGTTTCCAGCCATTCCTTCAACCCCTGATATCATGTGGCTGCA 7541  
QY 7417 GGGGGGTGATATGCTTCCCAACCCCAACCGGAGCTCCCGAGGAGCGGAGCCCTCGC 7476  
DB 7542 GGGGGGTGATATGCTTCCCAACCCCAACCGGAGCTCCCGAGGAGCGGAGCCCTCGC 7601  
QY 7477 TGGCCCCCAACAGCCTGGGACAGAGAGCCCAAGCCACTGTCTGTGCAATGACAGAC 7536  
DB 7602 TGGCCCCCAACAGCCTGGGAGAGAGAGCCCAAGCCACTGTCTGTGCAATGACAGAC 7661  
QY 7537 ACTTTCGACAGCGAGTGACTAGAAACAGGCGGGGGGGGGCGGGCGGTGERSPERLUG 7596  
DB 7662 ACTTTCGACAGCGAGTGACTAGAAACAGGCGGGGGGGGGCGGGCGG-----TG 7711  
QY 7597 TCAAGTCCAGAGGACAGAGGACGGCCCTGACAGAGGGGGGGGCTGCGCATCCGCC 7656  
DB 7712 TCAAGTCCAGAGGACAGAGGACGGCCCTGACAGAGGGGGGGGCTGCGCATCCGCC 7771  
QY 7657 AACCAAGAAAGAGCCCTTGAAGTCCGCTGCGCTCATCATCTGTCCGTCCAGAGCGG 7716  
DB 7772 AACCAAGAAAGAGCCCTTGAAGTCCGCTTGAAGTCCGCGCCCGGGGCTGCGCTCAGA 7831  
QY 7717 GCATCTTGCCTGTCTTAAAGCCTTAACTAAGACTCCCGCCCGGGGCTGCGCTCAGA 7776  
DB 7832 GCATCTTGCCTGTCTTAAAGCCTTAACTAAGACTCCCGCCCGGGGCTGCGCTCAGA 7891  
QY 7777 CCTTACTAGGGGAGTCTTAACTGTGTGCTCGGGAAGGGAGGGGAAGGGGCGGGGAGGGG 7836  
DB 7892 CCTTACTAGGGGAGTCTTAACTGTGTGCTCGGGAAGGGAGGGGAAGGGGCGGGGAGGGG 7951  
QY 7837 GCAAGGCAAGGCTGTGGAGGCAACAGAGCGGCGGAGGCGGAGGAGCCCAAGAGAG 7896  
DB 7952 GCAAGGCAAGGCTGTGGAGGCAACAGAGCGGCGGAGGCGGAGGAGCCCAAGAGAG 8011  
QY 7897 GATGACCAAGCACTTCCAGGCACTGCTCCCGAATGCAATTGGAAACCAAGTCTTAA 7956  
DB 8012 GATGACCAAGCACTTCCAGGCACTGCTCCCGAATGCAATTGGAAACCAAGTCTTAA 8071  
QY 7957 CTGAGCTCGAGGCCCCCGGCGCTCCTCCGCTCCCATTCGCGCTTGAAGCAG 8016  
DB 8072 CTGAGCTCGAGGCCCCCGGCGCTCCTCCGCTCCCATTCGCGCTTGAAGCAG 8131  
QY 8017 ATGGAAGGAGGCGCTGTCCAGGCCCCAGTGGCGTGGTCCGACCAAGACTGCCCCA 8076  
DB 8132 ATGGAAGGAGGCGCTGTCCAGGCCCCAGTGGCGTGGTCCGACCAAGACTGCCCCA 8191  
QY 8077 GCCAAGCAATTTGCTGAAACCAAGTCAAGGCGAGTGGGCGGACAAAGGCGCAGGTGCG 8136  
DB 8192 GCCAAGCAATTTGCTGAAACCAAGTCAAGGCGAGTGGGCGGACAAAGGCGCAGGTGCG 8251  
QY 8137 GCGTGGGGGAAACGATGCTCCGAGGACTGAGACTGTTTTTTTCAACATGCTTGGCGAG 8196  
DB 8252 GCGTGGGGGAAACGATGCTCCGAGGACTGAGACTGTTTTTTTCAACATGCTTGGCGAG 8311  
QY 8197 CGGTGGGAAGAAAGCAGATGTAATGATGTGTTTAAAGGATATATTTTGAATAC 8256  
DB 8312 CGGTGGGAAGAAAGCAGATGTAATGATGTGTTTAAAGGATATATTTTGAATAC 8371  
QY 8257 CTTCAATGAATTAATTCAAGATGTTTTTCAAGAAAGACTTAACCAAGTATTACTGCTCG 8316

Db 8372 CTTCAATGAATTAATTCAGANGTTTTAGCAAGAGACTTACCCAGTATTACTGCTGC 8431  
QY 8317 TGTGCTTTGATCTCTGCTTACCGTTCAAGAGCGGTGTGCAAGCGGACAGTGCATGCC 8376  
Db 8432 TGTGCTTTGATCTCTGCTTACCGTTCAAGAGCGGTGTGCAAGCGGACAGTGCATGCC 8491  
QY 8377 CATCACTCGCAGAGCAAGAGGCGGCGGAGCTGCTGTCACGCGCTGCTGCTCCTCC 8436  
Db 8492 CATCACTCGCAGAGCAAGAGGCGGCGGAGCTGCTGTCACGCGCTGCTGCTCCTCC 8551  
QY 8437 CCGTCCCTTCCTTGGGCGAAGAAATTCAGATGCGTATTCCTGTGCGCGCATTTGGCGAGG 8496  
Db 8552 CCGTCCCTTCCTTGGGCGAAGAAATTCAGATGCGTATTCCTGTGCGCGCATTTGGCGAGG 8611  
QY 8497 TGTGCTATTCCTGCTATTCACACGTCGCTTAATTAATAAAGCAATTAATCTCCAAA 8556  
Db 8612 TGTGCTATTCCTGCTATTCACACGTCGCTTAATTAATAAAGCAATTAATCTCCAAA 8671  
QY 8557 AAAAAAAAAAAAAA 8571  
Db 8672 AAAAAAAAAAAAAA 8686

RESULT 5  
US-10-174-014-4

; Sequence 4, Application US/10174014  
; Publication No. US20040005292A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Susan M. Freier  
; APPLICANT: Kenneth W. Doble  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION  
; FILE REFERENCE: PTS-0012  
; CURRENT APPLICATION NUMBER: US/10/174,014  
; CURRENT FILING DATE: 2002-06-17  
; NUMBER OF SEQ ID NOS: 73  
; SEQ ID NO 4  
; LENGTH: 8686  
; TYPE: DNA  
; ORGANISM: H. sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (157)...(7680)  
US-10-174-014-4

Query Match 97.2%; Score 8324.6; DB 6; Length 8686;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 8472; Conservative 1; Mismatches 34; Indels 88; Gaps 4;

QY 1 CATGTCGGGCTCCACACAGCTTGTCACACAGCGTGGAGGCGCACTGAGCCCGGCTACCC 60  
Db 156 CATGTCGGGCTCCACACAGCTTGTCACACAGCGTGGAGGCGCACTGAGCCCGGCTACCC 215  
QY 61 GCCCCACAGCCTTTCTTACCCAGTGCAGATCGCCCGGACGCAACAGGAGTGGCTCTCT 120  
Db 216 GCCCCACAGCCTTTCTTACCCAGTGCAGATCGCCCGGACGCAACAGGAGTGGCTCTCT 275  
QY 121 GAGATACCAAGCACTCCCGGACGATATGCTCCCACTGTCGGCGGCTCATATCA 180  
Db 276 GAGATACCAAGCACTCCCGGACGATATGCTCCCACTGTCGGCGGCTCATATCA 335  
QY 181 GCCCCAGCGGCGGAGGCGCTCCCTGCTGTAGTTCCAGCCCGGGAATGAACGTCCTCA 240  
Db 336 GCCCCAGCGGCGGAGGCGCTCCCTGCTGTAGTTCCAGCCCGGGAATGAACGTCCTCA 395  
QY 241 GAGGCTCCACCTGCGGCGCAAGTCCCATATACCTGCCCGAGCTGGGGAAGTCAAGAT 300  
Db 396 GAGGCTCCACCTGCGGCGCAAGTCCCATATACCTGCCCGAGCTGGGGAAGTCAAGAT 455  
QY 301 GAGGCTTCATTGAAGAGAGCGCCCTGCTTGAAGCTGCTGACCCCTGCTGACCC 360  
Db 456 GAGGCTTCATTGAAGAGAGCGCCCTGCTTGAAGCTGCTGACCCCTGCTGACCC 515

QY 361 GTCAACCCCTGTGCTGCGCAAGGCGCAGGCTGCGGATCTGAAGACTTCAACCAAGACCTGAG 420  
Db 516 GTCAACCCCTGTGCTGCGCAAGGCGCAGGCTGCGGATCTGAAGACTTCAACCAAGACCTGAG 575  
QY 421 CCGTACCGGCGCAAGCTGGAACCGGTGTCTCCCCCGGACCCCGGACACTGACCCCTGAGCT 480  
Db 576 CCGTACCGGCGCAAGCTGGAACCGGTGTCTCCCCCGGACCCCGGACACTGACCCCTGAGCT 635  
QY 481 GAGGCTGTGCGCGCACCGGCTGTCCAGAGAGAGCTGATCCAGAACATGACCCGCTGGA 540  
Db 636 GAGGCTGTGCGCGCACCGGCTGTCCAGAGAGAGCTGATCCAGAACATGACCCGCTGGA 695  
QY 541 CCGAGAGATCAACCATGTTAGAGCAGACAGATCTTAACTGAAGAAAGCAGCAACAGCT 600  
Db 696 CCGAGAGATCAACCATGTTAGAGCAGACAGATCTTAACTGAAAGAAAGCAGCAACAGCT 755  
QY 601 GAGAGAGAGGCTGCGCAAGCCCGCCGACCTGGAAGAGCCCGGTTCACCGCGCCATGGA 660  
Db 756 GAGAGAGAGGCTGCGCAAGCCCGCCGACCTGGAAGAGCCCGGTTCACCGCCATGGA 815  
QY 661 GTGGAAGCACCGCAGGCTGTGTGACAGATCATCTACAGAGAAACCGAGAAAGGCTGAAGC 720  
Db 816 GTGGAAGCACCGCAGGCTGTGTGACAGATCATCTACAGAGAAACCGAGAAAGGCTGAAGC 875  
QY 721 TGCAATCGGATTTGGAAGGCTTGGGCGCCAGGTGAGCTGCGCTGTACCAACGAGCC 780  
Db 876 TGCAATCGGATTTGGAAGGCTTGGGCGCCAGGTGAGCTGCGCTGTACCAACGAGCC 935  
QY 781 CTCGACACCCCGGCGAGTATCATGAGAAATCAATAAACAAGCGGAGTCCGGAAGAGCT 840  
Db 936 CTCGACACCCCGGCGAGTATCATGAGAAATCAATAAACAAGCGGAGTCCGGAAGAGCT 995  
QY 841 AATCTTACTTCAAGAGAGAGAAATCAGCTCGGAAACATGGAAGCAGAAAGTTCTGCA 900  
Db 996 AATCTTACTTCAAGAGAGAGAAATCAGCTCGGAAACATGGAAGCAGAAAGTTCTGCA 1055  
QY 901 GCGCTATGACCAAGCTCATGAGGCGCTTGGAAAAAAGGTGAGCGCATGAAAAACAACC 960  
Db 1056 GCGCTATGACCAAGCTCATGAGGCGCTTGGAAAAAAGGTGAGCGCATGAAAAACAACC 1115  
QY 961 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGGAGTCTAACGAAAGCAAGTCCCTGAGAT 1020  
Db 1116 CCGGCGCGCGGCGCAAGAGAGCAAGGTGCGGAGTCTAACGAAAGCAAGTCCCTGAGAT 1175  
QY 1021 CCGGAAGAGCGCGGAGCTGCAAGAGCGCATGCAAGGAGGCGGCGGAGCGGAGGAGTGG 1080  
Db 1176 CCGGAAGAGCGCGGAGCTGCAAGAGCGCATGCAAGGAGGCGGCGGAGCGGAGGAGTGG 1232  
QY 1081 GCTGTTCATGTCGCGCGCGCCGAGGAGCAAGAGGTGTCAAGATCATGATGCGCTCTC 1140  
Db 1233 GCTGTTCATGTCGCGCGCGCCGAGGAGCAAGAGGTGTCAAGATCATGATGCGCTCTC 1292  
QY 1141 AGAGCAGAGAACTTGAAGAAAGCAGATGCGGACGCTGAGTCCCGCCCATCTCTGA 1200  
Db 1293 AGAGCAGAGAACTTGAAGAAAGCAGATGCGGACGCTGAGTCCCGCCCATCTCTGA 1352  
QY 1201 GAGGCTGACCAAGCGGCGCATCAAGTTATCATCAACATGAACGAGGCTTATGCGCGAC 1260  
Db 1353 GAGGCTGACCAAGCGGCGCATCAAGTTATCATCAACATGAACGAGGCTTATGCGCGAC 1412  
QY 1261 GAAAGTGTAAAGAACCGCCAGGCTCATGAACATGTGAGTGAAGCAGAGAAAGAGACCTT 1320  
Db 1413 GAAAGTGTAAAGAACCGCCAGGCTCATGAACATGTGAGTGAAGCAGAGAAAGAGACCTT 1472  
QY 1321 CCGGGAAGAGTTTATGACAGATCCCAAGAACTTTTGGCTGATCGATCTTCTGAGAG 1380  
Db 1473 CCGGGAAGAGTTTATGACAGATCCCAAGAACTTTTGGCTGATCGATCTTCTGAGAG 1532  
QY 1381 GAAAGCAGTGTGAGTGTGCTCTATTATCACTGACTGAAGAAATGAGAACTATTA 1440  
Db 1533 GAAAGCAGTGTGAGTGTGCTCTATTATCACTGACTGAAGAAATGAGAACTATTA 1592



1441 GAGCTGTGTGAGACGAGCTATCGCGCCGCGCAGAGGCGAGCAACAGCAGCA 1500  
1593 GAGCCTGTGTGAGACGAGCTATCGCGCCGCGCAGAGGCGAGCAACAGCAGCA 1552  
1501 GCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1560  
1653 GCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1712  
1561 AGATGAG 1620  
1713 AGATGAG 1772  
1621 CGACAG 1680  
1773 CGACAG 1832  
1681 GAG 1740  
1833 GAG 1892  
1741 CGGCAATCACCCGCTCAATGAGTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800  
1893 CGGCAATCACCCGCTCAATGAGTAAATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1952  
1801 GAGCGCGAGAGCTGCTCCATGAGCTGAAATGAGAGTTCGCTGAGCAGAGAGAGAAAT 1860  
1953 GAGCGCGAGAGCTGCTCCATGAGCTGAAATGAGAGTTCGCTGAGCAGAGAGAGAAAT 2012  
1861 GAG 1920  
2013 GAG 2072  
1921 GGTGGGCTCCAAACGATGTCGAGAGTAAAGAACTTCTACTTCACTCAAGAGAGAGAG 1980  
2073 GGTGGGCTCCAAACGATGTCGAGAGTAAAGAACTTCTACTTCACTCAAGAGAGAGAG 2132  
1981 GAACTCTGATGAGATCTTGCAGAGAGCAACAGCTGAAATGAGAGAGAGAGAGAGAG 2040  
2133 GAACTCTGATGAGATCTTGCAGAGAGCAACAGCTGAAATGAGAGAGAGAGAGAGAG 2192  
2041 GAG 2100  
2193 GAG 2252  
2101 GAG 2160  
2253 GAG 2312  
2161 GGTGAGAGCTTATCATGCTCTGAGAAATGAGAGTGGCCAGAGAGAGAGAGAGAG 2220  
2313 GGTGAG-----AGC 2321  
2221 CACTGTCAACAAAGCTCAGAGACCGAGAGATCCCTCTCTCAACATGAGAGCGCCAA 2280  
2322 CACTGTCAACAAAGCTCAGAGACCGAGAGATCCCTCTCTCAACATGAGAGCGCCAA 2281  
2281 GAGACAG 2340  
2382 GAGACAG 2441  
2341 AGGCGCACCAACCCCAACAGAGAGAGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400  
2442 AGGCGCACCAACCCCAACAGAGAGAGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 2501  
2401 TGAAGCACCGAG 2460  
2502 TGAAGCACCGAG 2561  
2461 TGT 2520  
2562 TGT 2621  
2521 GAG 2580

2622 GAG 2681  
2581 GCGGCTCAAG 2640  
2682 GCGGCTCAAG 2741  
2641 GAG 2700  
2742 GAG 2801  
2701 CGGCAAG 2760  
2802 CGGCAAG 2861  
2761 CTGCAATGAG 2820  
2862 CTGCAATGAG 2921  
2821 AAGGCGCAGCTCTCAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
2922 AAGGCGCAGCTCTCAACCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2981  
2881 ACTGAGCTTGAAGCAG 2940  
2982 ACTGAGCTTGAAGCAG 3041  
2941 AGTCATGAG 3000  
3042 AGTCATGAG 3101  
3001 GCGACCGCAAAACCTGACAG 3060  
3102 GCGACCGCAAAACCTGACAG 3161  
3061 GGGCAAG 3100  
3162 GGGCAAG 3221  
3101 ----GCTTTCGAG 3156  
3222 CCGAG 3281  
3157 CCGAG 3216  
3282 CCGAG 3341  
3217 CCGAG 3276  
3342 CCGAG 3401  
3277 CCGAG 3336  
3402 CCGAG 3461  
3337 CAG 3396  
3462 CAG 3521  
3397 GCTCAAG 3456  
3522 GCTCAAG 3581  
3457 GCTCAAG 3516  
3582 GCTCAAG 3641  
3517 GCTCAAG 3576  
3642 GCTCAAG 3701  
3577 GCTCAAG 3636

Db 3702 GTCCGTGCTGAGAGGAGACAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCAACAAAGCAT 3761  
QY 3637 TCCAGACA CACGGGTGGCCCTCGAACA CCGGCATCATACCGCGGCTTCATCAACCCACGG 3696  
Db 3762 TCCAGACA CACGGGTGGCCCTCGAACA CCGGCATCATACCGCGGCTTCATCAACCCACGG 3821  
QY 3697 CACGCAAGTCCCGCAGACACCGTGTATACAAAGGACCATCAACAGGATCATTCGGAGAGACAGCCC 3756  
Db 3822 CACGCAAGTCCCGCAGACACCGTGTATACAAAGGACCATCAACAGGATCATTCGGAGAGACAGCCC 3881  
QY 3757 GAGTGCCTTGAACCGCGCCGGAGAGACAGCTGCGCCAAAGGGCCAGTCAATTAAGAAAG 3816  
Db 3882 GAGTGCCTTGAACCGCGCCGGAGAGACAGCTGCGCCAAAGGGCCAGTCAATTAAGAAAG 3941  
QY 3817 CAAGAAAGGCCCAGTCTTGTCTTATAGAGGTGGCATGTCTGTGACCCAGTCTCAAGGA 3876  
Db 3942 CAAGAAAGGCCCAGTCTTGTCTTATAGAGGTGGCATGTCTGTGACCCAGTCTCAAGGA 4001  
QY 3877 GGAACGGAGAAACACACTCAAGAACCCGCCATGAAACGGCCGCCCAAGCCGACCTATGA 3936  
Db 4002 GGAACGGAGAAACACACTCAAGAACCCGCCATGAAACGGCCGCCCAAGCCGACCTATGA 4061  
QY 3937 CATGATGAGAGGCGCGGTGGAGAGACCATCTCTCAAGCAGCATGAAAGTCTCATGGG 3996  
Db 4062 CATGATGAGAGGCGCGGTGGAGAGACCATCTCTCAAGCAGCATGAAAGTCTCATGGG 4121  
QY 4122 CGGTGCTCATCCCGCCGAGAGCACA CAGCCCCACCACTCAAAAGACAGCACCATCCG 4181  
Db 3997 CGGTGCTCATCCCGCCGAGAGCACA CAGCCCCACCACTCAAAAGACAGCACCATCCG 4056  
QY 4057 CGGTGCTCATCAACAAGGATATCCCTCGGTCTTACGTGAGAGAGACAGAGAACTTACTTGG 4116  
Db 4182 CGGTGCTCATCAACAAGGATATCCCTCGGTCTTACGTGAGAGAGACAGAGAACTTACTTGG 4241  
QY 4117 TCGGAGAGCCAAAGCTCTTAAAGCGGAGGGACA CGCTCCGCCCAACCGCCTCAAGGA 4176  
Db 4242 TCGGAGAGCCAAAGCTCTTAAAGCGGAGGGACA CGCTCCGCCCAACCGCCTCAAGGA 4301  
QY 4177 CCTGAACCGAGGCTTACAAAGACGAGGCCCTGGAGCCCCCTTGAAGCTGAAGCCGACCATGA 4236  
Db 4302 CCTGAACCGAGGCTTACAAAGACGAGGCCCTGGAGCCCCCTTGAAGCTGAAGCCGACCATGA 4361  
QY 4237 GGGGCTGTGTGGCAAGTGAAGAAGCGGGCCGCTTCATTCATGAGATTCGCCCGGAGGA 4296  
Db 4362 GGGGCTGTGTGGCAAGTGAAGAAGCGGGCCGCTTCATTCATGAGATTCGCCCGGAGGA 4421  
QY 4297 GCTGCGGACA CAGGCCGAGCTGCCCTGGCCCGCGGCGGCTCAAGAGAGGCTTCATCAC 4356  
Db 4422 GCTGCGGACA CAGGCCGAGCTGCCCTGGCCCGCGGCGGCTCAAGAGAGGCTTCATCAC 4481  
QY 4357 GCAAGGCA CCCCCGCTCAAGTACGACA CCGGCGGCTTCACACTGGCTCCAAAAGACAGA 4416  
Db 4482 GCAAGGCA CCCCCGCTCAAGTACGACA CCGGCGGCTTCACACTGGCTCCAAAAGACAGA 4541  
QY 4417 GCTAAGCTCCCTCATTCGAGAGCCCGGCGGAGAGGTTTCCCAACCTCGTGCACCCGCTGAATGT 4476  
Db 4542 GCTAAGCTCCCTCATTCGAGAGCCCGGCGGAGAGGTTTCCCAACCTCGTGCACCCGCTGAATGT 4601  
QY 4477 GATGCGCA CCGCCGAGCACTGGAAGTGTCTGCTACGAGAGAGGCTGAAGAGCGGCGC 4536  
Db 4602 GATGCGCA CCGCCGAGCACTGGAAGTGTCTGCTACGAGAGAGGCTGAAGAGCGGCGC 4661  
QY 4537 AAGGAACCGCAGAGCTCGGGGGGCTCAATTGCGGCGGCGGCCGCGTCAATGTGTCTGA 4596  
Db 4662 AAGGAACCGCAGAGCTCGGGGGGCTCAATTGCGGCGGCGGCCGCGTCAATGTGTCTGA 4721  
QY 4597 GCTGGGTAAAGCCCGGAGAGAGCCCTTGAACCTTATAGAGACACA CCGGAGCACTTTGGCCGG 4656  
Db 4722 GCTGGGTAAAGCCCGGAGAGAGCCCTTGAACCTTATAGAGACACA CCGGAGCACTTTGGCCGG 4781  
QY 4657 CCACTTCCACAGAGTTCGCGCTGTGACA TCGGAGAGCCACGCGCGCTGACAGAGGG 4716  
Db 4782 CCACTTCCACAGAGTTCGCGCTGTGACA TCGGAGAGCCACGCGCGCTGACAGAGGG 4841

QY 4717 CAGCCTTTGCTCAGCAAGGCATTCAGAGACGAAAGCTGAGCTGAGCGCTGTGAGAT 4776  
Db 4842 CAGCCTTTGCTCAGCAAGGCATTCAGAGACGAAAGCTGAGCTGAGCGCTGTGAGAT 4901  
QY 4777 GCGCAAGTCCCGCAGACACCGTGCAGAGACCAACCAACCCCATTCGSCCTATGA 4836  
Db 4902 GCGCAAGTCCCGCAGACACCGTGCAGAGACCAACCAACCCCATTCGSCCTATGA 4961  
QY 4837 GCACTGTCTTGGGGCGTGAAGTGGGTGAACCTGTATCGACAGCAATTCCTGGGCTT 4896  
Db 4962 GCACTGTCTTGGGGCGTGAAGTGGGTGAACCTGTATCGAGCCCAATTCCTGGGCTT 5021  
QY 4897 GCACTGTCTTGAATCCCGGAGATTCCTTGTGAACGAGCGGTGTGCTTACTTGGC 4956  
Db 5022 GCACTGTCTTGAATCCCGGAGATTCCTTGTGAACGAGCGGTGTGCTTACTTGGC 5081  
QY 4957 CCGACACCTGGCCCGCAACCCCACTTACCCGACCTGTATCCCACTTACTCATCCGCG 5016  
Db 5082 CCGACACCTGGCCCGCAACCCCACTTACCCGACCTGTATCCCACTTACTCATCCGCG 5141  
QY 5017 CTACCCCGACACGCGGCGCTGAGAAACCGGAGACATCATATGACTATCATCACTTC 5076  
Db 5142 CTACCCCGACACGCGGCGCTGAGAAACCGGAGACATCATATGACTATCATCACTTC 5201  
QY 5077 GCAAGATATGACCAACAACCGGCCACCTGCGCCAGCGAGCTGATATGCTGAAGGG 5136  
Db 5202 GCAAGATATGACCAACAACCGGCCACCTGCGCCAGCGAGCTGATATGCTGAAGGG 5261  
QY 5137 CCTCTGCGCCCGCGAGTCTGTGCTGAGCACTCAACTACGCTGGGGTCCCGAGGATCAT 5196  
Db 5262 CCTCTGCGCCCGCGAGTCTGTGCTGAGCACTCAACTACGCTGGGGTCCCGAGGATCAT 5321  
QY 5197 GCACTGTTCCAAGTGCACAACCTGCTGTGTCTGTGCCCCCGACACAAGCAACCCAGC 5256  
Db 5322 GCACTGTTCCAAGTGCACAACCTGCTGTGTCTGTGCCCCCGACACAAGCAACCCAGC 5381  
QY 5257 CACCGCATATGACCGGCTTTGCTTACCTCCCAACCGGCGCCGACCTTACGAGAGCGGCA 5316  
Db 5382 CACCGCATATGACCGGCTTTGCTTACCTCCCAACCGGCGCCGACCTTACGAGAGCGGCA 5441  
QY 5317 CAGGAGTCCCCCACTCTCCCGAGAGGTCCAAACA CATTGACAAAACCAACCAACGTC 5376  
Db 5442 CAGGAGTCCCCCACTCTCCCGAGAGGTCCAAACA CATTGACAAAACCAACCAACGTC 5501  
QY 5377 CTCTGTCCGAGCGGAGCGAGACCGGGAATGAGAGCGGAGACCGGGAATCGGAGACGGAAAA 5436  
Db 5502 CTCTGTCCGAGCGGAGCGAGACCGGGAATGAGAGCGGGAATCGGGAATCGGGAAGCGGAAAA 5561  
QY 5437 GTCCATCTCTACGTCACACGACGAGTGAAGACACGACCCCATTTGGAACCTGGTACGA 5496  
Db 5562 GTCCATCTCTACGTCACACGACGAGTGAAGACACGACCCCATTTGGAACCTGGTACGA 5621  
QY 5497 GCAAGAGCGGAGCAGACGAGGAGAGCGGGGAGTGGGGGACAGACAGCGGCGCGC 5556  
Db 5622 GCAAGAGCGGAGCAGACGAGGAGAGCGGGGAGTGGGGGACAGACAGCGGCGCGC 5681  
QY 5557 CTCCCACTTCCATGCGCAACGACACTCGCCCATCTTCCCTCGAACCCAGATGCCCTCA 5616  
Db 5682 CTCCCACTTCCATGCGCAACGACACTCGCCCATCTTCCCTCGAACCCAGATGCCCTCA 5741  
QY 5617 GCAAGAGCCCAAGTGTCTTACACAACAAGGATGAAGGGTATCATACCGCTGTGAGGC 5676  
Db 5742 GCAAGAGCCCAAGTGTCTTACACAACAAGGATGAAGGGTATCATACCGCTGTGAGGC 5801  
QY 5677 CAGCAAGCCCAAGTGTCTTACACAACAAGGATGAAGGGTATCATACCGCTGTGAGGC 5736  
Db 5802 CAGCAAGCCCAAGTGTCTTACACAACAAGGATGAAGGGTATCATACCGCTGTGAGGC 5861  
QY 5737 ATTCCCACTTGCACCACTGCGCACTGAGCGGACACCTCGATGGGGTCTACCTTACCT 5796  
Db 5862 ATTCCCACTTGCACCACTGCGCACTGAGCGGACACCTCGATGGGGTCTACCTTACCT 5921

QY 5797 CATGAGCCCGCTTGTGCTGCCAAGAGAGCCCGCCCGGCTGCCCGCCAGAGCGGCCCG 5856  
DB |||||  
QY 5922 CATGAGCCCGCTTGTGCTGCCAAGAGAGCCCGCCCGGCTGCCCGCCAGAGCGGCCCG 5981  
DB |||||  
QY 5857 AACGACACCGGCGCAATGCTTCTCTGCGCAAGCCCGCCGCTCCGGGCTGGAGCCCG 5916  
DB |||||  
QY 5982 AACGACACCGGCGCAATGCTTCTCTGCGCAAGCCCGCCGCTCCGGGCTGGAGCCCG 6041  
DB |||||  
QY 5917 CTCTCTCCCGCAAGAGGCTCGGAGCCCGCCCTTATGTCCTCTGTCTGGACACG 5976  
DB CTCTCTCCCGCAAGAGGCTCGGAGCCCGCCCTTATGTCCTCTGTCTGGACACG 6101  
QY 5977 CACCATGCGCCGCAACCCCTGCGAAGAACTTCGACCTCAACAGCGACCCCGAACCCG 6036  
DB |||||  
QY 6102 CACCATGCGCCGCAACCCCTGCGAAGAACTTCGACCTCAACAGCGACCCCGAACCCG 6161  
DB |||||  
QY 6037 GGGCGCACTGCTGCTGGGCTTCGGACCCCGCAACCGGAAAGAACTCAAAAGTAAACCTTTTC 6096  
DB GGGCGCACTGCTGCTGGGCTTCGGACCCCGCAACCGGAAAGAACTCAAAAGTAAACCTTTTC 6221  
QY 6097 CATCCAGAACTGGAATCCGTTCTTGGGTTTACACGCGACAGGCTTACAGCCCGGAAG 6156  
DB CATCCAGAACTGGAATCCGTTCTTGGGTTTACACGCGACAGGCTTACAGCCCGGAAG 6281  
QY 6157 GGTGAGCCCGCTCAAGCCCTGTGAGCTCACCCAGTCTGACCAAGAGGGGCTCCCAA 6216  
DB GGTGAGCCCGCTCAAGCCCTGTGAGCTCACCCAGTCTGACCAAGAGGGGCTCCCAA 6341  
QY 6217 GCACTCTGGAAGAGCTTGAACAAAGGCACTTGGAGGGGAGCTGCGGCCCAAGACCAAG 6276  
DB GCACTCTGGAAGAGCTTGAACAAAGGCACTTGGAGGGGAGCTGCGGCCCAAGACCAAG 6401  
QY 6342 GCACTCTGGAAGAGCTTGAACAAAGGCACTTGGAGGGGAGCTGCGGCCCAAGACCAAG 6401  
DB |||||  
QY 6277 CCCCGGAAGCTTGGGCGGAGGCGCCACCTCCCACTGCGGCGCTGCTGAGAG 6336  
DB CCCCGGAAGCTTGGGCGGAGGCGCCACCTCCCACTGCGGCGCTGCTGAGAG 6461  
QY 6337 CCAGCCCTTGTCAAGCCCGCTGTGAGACCGGCCAGGGGTTCAAGAGTCAACAGCGGCT 6396  
DB CCAGCCCTTGTCAAGCCCGCTGTGAGACCGGCCAGGGGTTCAAGAGTCAACAGCGGCT 6521  
QY 6397 GGTCACTCTGCGCCAGCAATCATGTAGGTCAATCAACAGAACTAACCCCGGCAACACCC 6456  
DB GGTCACTCTGCGCCAGCAATCATGTAGGTCAATCAACAGAACTAACCCCGGCAACACCC 6581  
QY 6522 GGTCACTCTGCGCCAGCAATCATGTAGGTCAATCAACAGAACTAACCCCGGCAACACCC 6581  
DB |||||  
QY 6457 AACGAGAGCTCAGGAGCAACCCCTGCGGCGCCCTCTTACTGCTCCCTGGGAGCAGCTGCC 6516  
DB AACGAGAGCTCAGGAGCAACCCCTGCGGCGCCCTCTTACTGCTCCCTGGGAGCAGCTGCC 6641  
QY 6582 AACGAGAGCTCAGGAGCAACCCCTGCGGCGCCCTCTTACTGCTCCCTGGGAGCAGCTGCC 6641  
DB |||||  
QY 6517 GGTCTGTGACCTTCGCGCGCCCAACCAAGTACCTTACCTTCGCGCCCGGACCATGCTGC 6576  
DB GGTCTGTGACCTTCGCGCGCCCAACCAAGTACCTTACCTTCGCGCCCGGACCATGCTGC 6701  
QY 6577 CCCGCGCCCTGCTGCCCAAGGAGGCGCAAGAGGCTTCCAGAGCCAAACAAAGC 6636  
DB CCCGCGCCCTGCTGCCCAAGGAGGCGCAAGAGGCTTCCAGAGCCAAACAAAGC 6761  
QY 6702 CCCGCGCCCTGCTGCCCAAGGAGGCGCAAGAGGCTTCCAGAGCCAAACAAAGC 6761  
DB |||||  
QY 6637 GTTGGGCTTGGGAGGAGGAGGAGGATTTGAACCTGTGCCCAACCGGAGGAGCAAGC 6696  
DB GTTGGGCTTGGGAGGAGGAGGAGGATTTGAACCTGTGCCCAACCGGAGGAGCAAGC 6821  
QY 6697 GAGGCGAGGAGCACTCCGAGAGTGTGTATCCGCTGTGTACCGGAGTGGGAAACAGC 6756  
DB GAGGCGAGGAGCACTCCGAGAGTGTGTATCCGCTGTGTACCGGAGTGGGAAACAGC 6881  
QY 6822 GAGGCGAGGAGCACTCCGAGAGTGTGTATCCGCTGTGTACCGGAGTGGGAAACAGC 6881  
DB |||||  
QY 6757 GAGGCGAGGAGTGGGCTTCAAGTCTTCAAGGCAACCAAGCAGCGCCAGCTTCTTT 6816  
DB |||||  
QY 6882 GAGGCGAGGAGTGGGCTTCAAGTCTTCAAGGCAACCAAGCAGCGCCAGCTTCTTT 6941  
DB |||||  
QY 6817 CAGCAAGCTGACCGAGAGCACTCCGCAATGGTCAAGTCCAAAGACAAAGATCAACA 6876  
DB CAGCAAGCTGACCGAGAGCACTCCGCAATGGTCAAGTCCAAAGACAAAGATCAACA 7001  
QY 6942 CAGCAAGCTGACCGAGAGCACTCCGCAATGGTCAAGTCCAAAGACAAAGATCAACA 7001  
DB |||||  
QY 6877 GAAGCTGAACACCCAAACCGGAATGAGCCTGAATTAATAATACGCAAGCCTGGAGCGA 6936

DB |||||  
QY 7002 GAAGCTGAACACCCAAACCGGAATGAGCCTGAATTAATAATACGCAAGCCTGGAGCGA 7061  
DB |||||  
QY 6937 GATCTTCAATATATGCGCCCATCAACGGAACAGGCTTATGACTTATAGAACCAAGCGCT 6996  
DB GATCTTCAATATATGCGCCCATCAACGGAACAGGCTTATGACTTATAGAACCAAGCGCT 7121  
QY 6997 GCAGGAACATGCGCAGCAACCAATGAGGCTGAGAGGCTATATTAAGAAAGCACTCATGAG 7056  
DB GCAGGAACATGCGCAGCAACCAATGAGGCTGAGAGGCTATATTAAGAAAGCACTCATGAG 7181  
QY 7057 TAAATATGACAGTGGGAAAGTATCCCGCGCTCAGCGCAATGCTTTTAAACCTTTGA 7116  
DB TAAATATGACAGTGGGAAAGTATCCCGCGCTCAGCGCAATGCTTTTAAACCTTTGA 7241  
QY 7117 TGCAGTGCAGGCTGCGCGCTGTATGCTTAAACCGCTGTGACGAGAGTACCA 7176  
DB TGCAGTGCAGGCTGCGCGCTGTATGCTTAAACCGCTGTGACGAGAGTACCA 7301  
QY 7242 TGCAGTGCAGGCTGCGCGCTGTATGCTTAAACCGCTGTGACGAGAGTACCA 7301  
DB |||||  
QY 7177 CACACTCACTTGCAGAGTGTGAGCGCGGGAAGGCAAGTCTTGGCAGACCCAGACCGG 7236  
DB CACACTCACTTGCAGAGTGTGAGCGCGGGAAGGCAAGGCTTGGCAGACCCAGACCGG 7361  
QY 7237 AAAAGCAAGTCCCGGCGCTGCGGCTTGGCATCTGGGAGACCGGCAACCTTGTCTCTTC 7296  
DB AAAAGCAAGTCCCGGCGCTGCGGCGCTTGGCATCTGGGAGACCGGCAACCTTGTCTCTTC 7421  
QY 7362 AAAAGCAAGTCCCGGCGCTGCGGCGCTTGGCATCTGGGAGACCGGCAACCTTGTCTCTTC 7421  
DB |||||  
QY 7297 AGTGCACCTGAGGAGAGACTGCAACCGCGGAGCGCGCTCACCAACCGGCTGAGGAGGA 7356  
DB AGTGCACCTGAGGAGAGACTGCAACCGCGGAGCGCGCTCACCAACCGGCTGAGGAGGA 7481  
QY 7422 AGTGCACCTGAGGAGAGACTGCAACCGCGGAGCGCGCTCACCAACCGGCTGAGGAGGA 7481  
DB |||||  
QY 7357 CAGGCGCTTGTGCGGAGGTTTCCAGCGCAATCCCTTACCAACCCCTGATCATGCGCTGCA 7416  
DB CAGGCGCTTGTGCGGAGGTTTCCAGCGCAATCCCTTACCAACCCCTGATCATGCGCTGCA 7541  
QY 7417 GCGCGGCTGATGAGGCTTCCCAACCCCGGCGCTCCCGGCGGAGCGGCGCTTCCG 7476  
DB GCGCGGCTGATGAGGCTTCCCAACCCCGGCGCTCCCGGCGGAGCGGCGCTTCCG 7601  
QY 7542 GCGCGGCTGATGAGGCTTCCCAACCCCGGCGCTCCCGGCGGAGCGGCGCTTCCG 7601  
DB |||||  
QY 7477 TGGCCCCCAACAGCGCTGGGAGGAGGAGGCGCAAGGCACTGCTGTGCGAGTACAGAGC 7536  
DB TGGCCCCCAACAGCGCTGGGAGGAGGAGGCGCAAGGCACTGCTGTGCGAGTACAGAGC 7661  
QY 7602 TGGCCCCCAACAGCGCTGGGAGGAGGAGGCGCAAGGCACTGCTGTGCGAGTACAGAGC 7661  
DB |||||  
QY 7537 ACTCTCCAGCAGCGAGTGACTCAGAAACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 7596  
DB ACTCTCCAGCAGCGAGTGACTCAGAAACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7711  
QY 7597 TCAAGTCCAGCGAGGCAACAGGAGAGGCTTCAAGAGAGCGGCTTCCGACTTCCG 7656  
DB TCAAGTCCAGCGAGGCAACAGGAGAGGCTTCAAGAGAGCGGCTTCCGACTTCCG 7771  
QY 7712 TCAAGTCCAGCGAGGCAACAGGAGAGGCTTCAAGAGAGCGGCTTCCGACTTCCG 7771  
DB |||||  
QY 7657 AACCAAGGAAGAGGCGCCGTAATCCGCGCTGCGGCTCCCAATCTGCTCCGTAAGAGCGG 7716  
DB AACCAAGGAAGAGGCGCCGTAATCCGCGCTGCGGCTCCCAATCTGCTCCGTAAGAGCGG 7831  
QY 7772 AACCAAGGAAGAGGCGCCGTAATCCGCGCTGCGGCTCCCAATCTGCTCCGTAAGAGCGG 7831  
DB |||||  
QY 7717 GCATCTTGCCTGTCTTAAAGCCTTAACTTAAAGTCCCGCGCGGCTGAGCTTGTGACA 7776  
DB GCATCTTGCCTGTCTTAAAGCCTTAACTTAAAGTCCCGCGCGGCTGAGCTTGTGACA 7891  
QY 7777 CCTTACTCAGGAGATGTTTAACTTGTGCTCGGAAAGGAGGAGGAGGAGGAGGAGGAGG 7836  
DB CCTTACTCAGGAGATGTTTAACTTGTGCTCGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 7951  
QY 7892 CCTTACTCAGGAGATGTTTAACTTGTGCTCGGAAAGGAGGAGGAGGAGGAGGAGGAGG 7951  
DB |||||  
QY 7837 GGAAGCGAGGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7896  
DB GGAAGCGAGGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8011  
QY 7952 GGAAGCGAGGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8011  
DB |||||  
QY 7897 GATGACCAAGCACTTCAACGCACTGCTCCCGGATATGCAATTTGGAACCAAGTCTTAA 7956  
DB GATGACCAAGCACTTCAACGCACTGCTCCCGGATATGCAATTTGGAACCAAGTCTTAA 8071  
QY 8012 GATGACCAAGCACTTCAACGCACTGCTCCCGGATATGCAATTTGGAACCAAGTCTTAA 8071  
DB |||||  
QY 7957 CTGAGCTGCAAGCCCGCGGCTTCCCTTCCGCTTCCATCCGCTTAAAGGCTTGGACAG 8016

Db 8072 CTGAGCTGCAAGCCCCCGGCGCTCCCTCCGCTCCATCCCGCTTAAAGCGCTGTGAACAG 8131  
QY 8017 ATGAGCAGAGGCGCTGTCCAGAGCCCCAGTGCCTGTTCGGGTCCCAACAAGTCCCA 8076  
Db 8132 ATGAGCAGAGGCGCTGTCCAGAGCCCCAGTGCCTGTTCGGGTCCCAACAAGTCCCA 8191  
QY 8077 GCCAAGAGATTGTGGAAGAACCAAGTCAAGGCGAGTGGCGGACAAAGGCGAGGTGG 8136  
Db 8192 GCCAAGAGATTGTGGAAGAACCAAGTCAAGGCGAGTGGCGGACAAAGGCGAGGTGG 8251  
QY 8137 GCGTGGGGGGAACGAGATGCTCCGAGAGAGCTGATTTTTTTTCAACAATCGTGGCGAG 8196  
Db 8252 GCGTGGGGGGAACGAGATGCTCCGAGAGAGCTGATTTTTTTTCAACAATCGTGGCGAG 8311  
QY 8197 CGGTGGGAAGGAAGGAGAGTGAATGATGTGTGTGTTTACAGGGTATATTTTGTATAC 8256  
Db 8312 CGGTGGGAAGGAAGGAGAGTGAATGATGTGTGTGTTTACAGGGTATATTTTGTATAC 8371  
QY 8257 CTTCAATGAATTAATTCAGATGTTTTTACGCAAGAGAGCTTACCCAGTATTACTGTGC 8316  
Db 8372 CTTCAATGAATTAATTCAGATGTTTTTACGCAAGAGAGCTTACCCAGTATTACTGTGC 8431  
QY 8317 TGTGCTTTTGAATCTGTGCTTACCGTTCAAGAGCGTGGCAGGCGAGAGTGGTGAACC 8376  
Db 8432 TGTGCTTTTGAATCTGTGCTTACCGTTCAAGAGCGTGGCAGGCGAGAGTGGTGAACC 8491  
QY 8377 CATCATCGCAGAGCAAGAGGGCGGGAGCTGCTGTCAAGCGCCGCTGTCTCCCTC 8436  
Db 8492 CATCATCGCAGAGCAAGAGGGCGGGAGCTGCTGTCAAGCGCCGCTGTCTCCCTC 8551  
QY 8437 CCTCCCTTCTTGGGCAAGATGAATTCGATGCGTATCTGTGGCGGCAATTTGGCGAGG 8496  
Db 8552 CCTCCCTTCTTGGGCAAGATGAATTCGATGCGTATCTGTGGCGGCAATTTGGCGAGG 8611  
QY 8457 TGTGCTTTTGAATCTGTGCTTACCGTTCAAGAGCGTGGCAGGCGAGAGTGGTGAACC 8556  
Db 8612 TGTGCTTTTGAATCTGTGCTTACCGTTCAAGAGCGTGGCAGGCGAGAGTGGTGAACC 8671  
QY 8557 AAAAAAAAAAAAAA 8571  
Db 8672 AAAAAAAAAAAAAA 8686

RESULT 6  
US-10-723-860-1739  
; Sequence 1739, Application US/10723860  
; Publication No. US20040253606A1  
; GENERAL INFORMATION:  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Gineburg, Wendy M.  
; APPLICANT: Zlotnick, Albert  
; TITLE OF INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions &  
; FILE REFERENCE: 05882, 0193, NPO501  
; CURRENT APPLICATION NUMBER: US/10/723, 860  
; PRIOR FILING DATE: 2003-11-26  
; PRIOR APPLICATION NUMBER: 60/429, 739  
; NUMBER OF SEQ ID NOS: 8393  
; SOFTWARE: Patent version 3.2  
; SEQ ID NO 1739  
; LENGTH: 8686  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-723-860-1739

Query Match 97.2%; Score 8324.6; DB 8; Length 8686;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 8472; Conservative 1; Mismatches 34; Indels 88; Gaps 4;

QY 1 CATGTGCGGCTTCCACACAGCTTGTGGCAGACGTTGAGAGGCGCACTGAGCCCGCTTACCC 60  
Db 156 CATGTGCGGCTTCCACACAGCTTGTGGCAGACGTTGAGAGGCGCACTGAGCCCGCTTACCC 215

QY 61 GCCCCAGAGCTTTTCTTACCCAGTGCAGATGCCCCGAGCGACACGAGCTGGGCTCTT 120  
Db 216 GCCCCAGAGCTTTTCTTACCCAGTGCAGATGCCCCGAGCGACACGAGCTGGGCTCTT 275  
QY 121 GAGATACCAAGACCACTCCCGGAGCTATGCTTCCCACTGTTCGCGGGCTCATCATCA 180  
Db 276 GAGATACCAAGACCACTCCCGGAGCTATGCTTCCCACTGTTCGCGGGCTCATCATCA 335  
QY 181 GCCCCAGAGCGAGAGGCTCTCTGCTGTGTGAGTTCCAGCCCGGGAATGAACGATGCCA 240  
Db 336 GCCCCAGAGCGAGAGGCTCTCTGCTGTGTGAGTTCCAGCCCGGGAATGAACGATGCCA 395  
QY 241 GAGATCCACTGCGGCGCAGAGTCCACTCATACCTGCGGAGCTGGGGAATGCAGAT 300  
Db 396 GAGATCCACTGCGGCGCAGAGTCCACTCATACCTGCGGAGCTGGGGAATGCAGAT 455  
QY 301 GAGATTCATTAAGAACAGCGCCCTCGGCTTGAAGTGTGTGCTGACCCCGCTGGCGACC 360  
Db 456 GAGATTCATTAAGAACAGCGCCCTCGGCTTGAAGTGTGTGCTGACCCCGCTGGCGACC 515  
QY 361 GTACACCCCTGTGGCCAGCGGCGCAGCTGCGGAGTCTGAAGACCTCACCAAGAGCCGTAG 420  
Db 516 GTACACCCCTGTGGCCAGCGGCGCAGCTGCGGAGTCTGAAGACCTCACCAAGAGCCGTAG 575  
QY 421 CTTGACGCGGCAAGCTGGAACCGGTGTCTCCCTCCAGCCCTCCGCACTGACCTGTAGCT 480  
Db 576 CTTGACGCGGCAAGCTGGAACCGGTGTCTCCCTCCAGCCCTCCGCACTGACCTGTAGCT 635  
QY 481 GAGCTGTGTGCGGCGCAGCGGCTGTCCAGAGAGAGTGTATCCAGAACATGAGCCGGTGA 540  
Db 636 GAGCTGTGTGCGGCGCAGCGGCTGTCCAGAGAGAGTGTATCCAGAACATGAGCCGGTGA 695  
QY 541 CCGAGAGATCAACATGTGTAGAGCAGACAGATCTTAAGCTGAAGAGAGAGCAACACT 600  
Db 696 CCGAGAGATCAACATGTGTAGAGCAGACAGATCTTAAGCTGAAGAGAGAGCAACACT 755  
QY 601 GAGAGAGAGGCTGCGCAAGCGCCCGAGCTGTGAGAGCCCGTGTCAAGCCCGCATGGA 660  
Db 756 GAGAGAGAGGCTGCGCAAGCGCCCGAGCTGTGAGAGCCCGTGTCAAGCCCGCATGGA 815  
QY 661 GTCCAGACCGCGAGCGCTGTGTGAGATCTTACAGAGAGAACCGGAGAGAGGCTGAAGC 720  
Db 816 GTCCAGACCGCGAGCGCTGTGTGAGATCTTACAGAGAGAACCGGAGAGAGGCTGAAGC 875  
QY 721 TGACATGAGATTCGGAAGGCTGGGGGCGCAGGTGAGGCTGCGGTCAACACAGCC 780  
Db 876 TGACATGAGATTCGGAAGGCTGGGGGCGCAGGTGAGGCTGCGGTCAACACAGCC 935  
QY 781 CTCGACACCGCGCAGTATCTGAGAACATCAAAATTAACCAAGGCGATGCGAAGAGCT 840  
Db 936 CTCGACACCGCGCAGTATCTGAGAACATCAAAATTAACCAAGGCGATGCGAAGAGCT 995  
QY 841 AATCTTGTACTTCAAGAGAGAGATCAGCTTCGGAACATATGAGAGAGAGAGTTCGCA 900  
Db 996 AATCTTGTACTTCAAGAGAGAGATCAGCTTCGGAACATATGAGAGAGAGAGTTCGCA 1055  
QY 901 GCGCTATGACAGCTCATGAGAGGCTTGGAAAAAAGGTGAGGCGCATGCAAAACACCC 960  
Db 1056 GCGCTATGACAGCTCATGAGAGGCTTGGAAAAAAGGTGAGGCGCATGCAAAACACCC 1115  
QY 961 GCGCGCGCGGCGCAAGAGAGAGAGTTCGCGAGTATCTACAAAAAGCATTCCTGAGAT 1020  
Db 1116 GCGCGCGCGGCGCAAGAGAGAGAGTTCGCGAGTATCTACAAAAAGCATTCCTGAGAT 1175  
QY 1021 CCGCAAGAGCGCGAGCTGCAAGAGCGCATGCAAGAGGCTGGCGCGAGCGGAGCGTGG 1080  
Db 1176 CCGCAAGAGCGCGAGCTGCAAGAGCGCATGCAAGAGGCTGGCGCGAGCGGAGCGTGG 1232  
QY 1081 GCTGTTCATGTTCGCGCGCGCGCGAGAGACAGAGTGTCAAGATCATGATGCGCTGTC 1140  
Db 1233 GCTGTTCATGTTCGCGCGCGCGCGAGAGACAGAGTGTCAAGATCATGATGCGCTGTC 1292

QY	1141	AGAGCAGGAAACCTCGGAGAAAGACAGATGCGGCACGTGGCCGTGATCCGCCCATGCTGTA	1200
Db	1293	AGAGCAGGAAACCTCGGAGAAAGACAGATGCGGCACGTGGCCGTGATCCGCCCATGCTGTA	1352
QY	1201	CGAGCGCTACCAAGCGGACATCAAGTTGATCAACATGAAACGGGGCTTAATGAGCCGAC	1260
Db	1353	CGAGCGCTACCAAGCGGACATCAAGTTGATCAACATGAAACGGGGCTTAATGAGCCGAC	1412
QY	1261	GAAGGTGTACAAAGACCGCCAGGTCATGAAACATGTGAGTGAAGCAGAGAAAGACCTT	1320
Db	1413	GAAGGTGTACAAAGACCGCCAGGTCATGAAACATGTGAGTGAAGCAGAGAAAGACCTT	1472
QY	1321	CCGGGAGAAAGTTCAATGACAGCATCCCAAGAACTTTGAGCTGATTCGATATTTCTTGAGAG	1380
Db	1473	CCGGGAGAAAGTTCAATGACAGCATCCCAAGAACTTTGAGCTGATTCGATATTTCTTGAGAG	1532
QY	1381	GAAAGCAGTGGCTGAGTGGCTGCTCTTATTAATCACTGTACTAAGAAATGAGAACTTATA	1440
Db	1533	GAAAGCAGTGGCTGAGTGGCTGCTCTTATTAATCACTGTACTAAGAAATGAGAACTTATA	1592
QY	1441	GAGCCTGTGAGACGAGACCTATTCGCGCGCGGCAAGACCAAGCAGACAAACAGACACA	1500
Db	1593	GAGCCTGTGAGACGAGACCTATTCGCGCGCGGCAAGACCAAGCAGACAAACAGACACA	1652
QY	1501	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1560
Db	1653	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1712
QY	1561	AGATGAGAAAGAAAGAAAGAAAGAGGCGGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	1620
Db	1713	AGATGAGAAAGAAAGAAAGAAAGAGGCGGAAAGAAAGAAAGAAAGAAAGAAAGAAAG	1772
QY	1621	CGACAAAGAAAGACCTCTCAAGAGAAAGACAGACGACCTCAGGGGAGAGCAACGACGA	1680
Db	1773	CGACAAAGAAAGACCTCTCTCAAGAGAAAGACAGACGACCTCAGGGGAGAGCAACGACGA	1832
QY	1681	GAAAGAGGCTGTGGCTTCGCTCAAAAGGCGGCAAACTGCGCAACAGCCAGGAAAGCGCAAGG	1740
Db	1833	GAAAGAGGCTGTGGCTTCGCTCAAAAGGCGGCAAACTGCGCAACAGCCAGGAAAGCGCAAGG	1892
QY	1741	CCGCATCAACCCGCTCAATGCTTAATGAGGCCAACAAGGAGGAGGATCAACCCGCCAGCA	1800
Db	1893	CCGCATCAACCCGCTCAATGCTTAATGAGGCCAACAAGGAGGAGGATCAACCCGCCAGCA	1952
QY	1801	GAGCGCCAGGCTGGCTCATGAGAGCTGAATGAGAGTTCTGCTGAGACAGAAAGAAAT	1860
Db	1953	GAGCGCCAGGCTGGCTCATGAGAGCTGAATGAGAGTTCTGCTGAGACAGAAAGAAAT	2012
QY	1861	GGAAGACAGCCAAAGAAAGGTCTCTTGAAACACGCGCGCAACTGTGCGCATGCGCCGAT	1920
Db	2013	GGAAGACAGCCAAAGAAAGGTCTCTTGAAACACGCGCGCAACTGTGCGCATGCGCCGAT	2072
QY	1921	GGTGGGCTCCAAAGCTGTGTGCGAGTGTAAAGAACTTTCACTTCAACTAACAAAGAAAGGCA	1980
Db	2073	GGTGGGCTCCAAAGCTGTGTGCGAGTGTAAAGAACTTTCACTTCAACTAACAAAGAAAGGCA	2132
QY	1981	GAACTCTGATGAGATCTTTGAGACAGCAACAAGCTGAAGATGAGAGAGGAGAAAGCGCG	2040
Db	2133	GAACTCTGATGAGATCTTTGAGACAGCAACAAGCTGAAGATGAGAGAGGAGAAAGCGCG	2192
QY	2041	GAGGAAAGAAAGAAAGGCGCGGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2100
Db	2193	GAGGAAAGAAAGAAAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2252
QY	2101	GGAGGATGAGAGATGAGAGCGGTGGGCGGTGAGCGGAAATGAGAGAGGAGAGATGTGAGAGA	2160
Db	2253	GGAGGATGAGAGATGAGAGCGGTGGGCGGTGAGCGGAAATGAGAGAGGAGAGATGTGAGAGA	2312
QY	2161	GGCTTGAAGCCTTATCATGCTCTGTGGGAAATGAGAGTCCAGAGGGGAAATGCAATGCGCCAGC	2220
Db	2313	GGCTTGAAGCCTTATCATGCTCTGTGGGAAATGAGAGTCCAGAGGGGAAATGCAATGCGCCAGC	2320
QY	2221	CACGTCAACACAGCTCAGACACCGAGAGCATCCCTCTCTCACTGAGGCGGCA	2280

Db	2222	CACGTGTAAACAAGCTTCAAGACACCGAAGCATCCCTCTCTCAACTGAGCGGCGCAA	2381
QY	2281	GGACACAGGGGAGAAATGGGCCCCAAGCCCCCAGCCAGCCCTGGGCGCGACGAGGCCACCCCC	2340
Db	2382	GGACACAGGGGAGAAATGGGCCCCAAGCCCCCAGCCAGCCCTGGGCGCGCGAGCGGGCCACCCCC	2441
QY	2341	AGGCCCAACCCACCCCAACAACGAGAGACATCCCGGGCCCCCAATTGAGCCCACTCCGGGCTC	2400
Db	2442	AGGGCCACCCCAACCCCAACAACGAGAGACATCCCGGGCCCCCAATTGAGTCACACCCGGGCTC	2501
QY	2401	TGAAGCCACCGGAGACCCCTTACGGCCCCCAACAGACACCCCAATCGGCTCTGTGACCTCCCTCC	2460
Db	2502	TGAAGCCACCTTAAAGCCCCCTTACGGCCCCCAACAGACACCCCAATTCCTCTTCACTTCCTCC	2561
QY	2461	TGTGTGTCCCCCAAGAGAGAAAGAGAGAGACCGACAGACAGCGCCCCCAAGTGGAGAGAGGG	2520
Db	2562	TGTGTGTCCCCCAAGAGAGAAAGAGAGAGACCGACAGACAGCGCCCCCAAGTGGAGAGAGGG	2621
QY	2521	GGAGAGACAGAAAGCCCCCGCGGCTGAGAGACTGGACAGTGGACACAGAGAAAGGCTCGAGGA	2580
Db	2622	GGAGAGACAGAAAGCCCCCGCGGCTGAGAGACTGGACAGTGGACACAGAGAAAGGCTCGAGGA	2681
QY	2581	GCCCGTCAAGAGCGAAGTGTCAACGAGGAAGCCGAGAGAGGGGCGGGCCAAAGGGCAAGACGC	2640
Db	2682	GCCCGTCAAGAGCGAAGTGTCAACGAGGAAGCCGAGAGAGGGGCGGGCCAAAGGGCAAGACGC	2741
QY	2641	GGAGGCGCGTGAAGGCAACGCGCGGAGGGGCGCTCAAGGCAAGAAAGAAAGGAGGCGGAG	2700
Db	2742	GGAGGCGCGTGAAGGCAACGCGCGGAGGGGCGCTCAAGGCAAGAAAGAAAGGAGGCGGAG	2801
QY	2701	CGGACGAGGCGCAACACTGCCAAGAGCTCGGGCGCCCCCAAGACAGCGACTTCAGTGTAC	2760
Db	2802	CGGACGAGGCGCAACAACGCAAGAGCTCGGGGCCCCCAAGACAGCGACTTCAGTGTAC	2861
QY	2761	CTGCAATGTCAAGCGAGGTGATGAGCGCCGAGAGGGCGGCGACAGAAACCGGCTGTCTCTCC	2820
Db	2862	CTGCAATGTCAAGCGAGGTGATGAGCGCGGAGGGCGGCGACAGAAACCGGCTGTCTCTCC	2921
QY	2821	AAAGGCCAAGGCTCTTCAACCCCGAATGGCGACCCCCGGGGCCAAATGCTCAACCCAGAAAGCC	2880
Db	2922	AAAGGCCAAGGCTCTTCAACCCCGAATGGCGACCCCCGGGGCCAAATGCTCTCAACCCAGAAAGCC	2981
QY	2881	ACTGGAAGCTGAAAGCAGTGAAGCAGACGAGCGGCTCCATCCCCCCTTCCAGTCAACCA	2940
Db	2982	ACTGGAAGCTTGAAGCAGTGAAGCAGACGAGCGGCTCCATCCCCCCTTCCAGTCAACCA	3041
QY	2941	AGTCCATGAGCCCCCGGGAGGAGCGAGCTCCCAACAAAGCAGCTCCCCAGGCCCAAC	3000
Db	3042	AGTCCATGAGCCCCCGGGAGGAGCGAGCTCCCAACAAAGCAGCTCCCCAGGCCCAAC	3101
QY	3001	GGCAGCGGCAAAACCTTGACGCGGAGACGAGACGCCCCCAGAGGCTTGGACAGACGCCCG	3060
Db	3102	GGCAGCGGCAAAACCTTGACGCGGAGACGAGACGCCCCCAGAGGCTTGGACAGACGCCCG	3161
QY	3061	GGGCAAGAGAGAGAGCCCGGCAACCCCGCGCGACAAAGAG-----3100	
Db	3162	GGGCAAGAGAGAGAGCCCGGCAACCCCGCGCGCAAAAGAGGCTGTGTTCTT	3221
QY	3101	-----GCCTTGCGACGCGAGGCCCAAGAGCTGTGGGAAACCCCTTGTGTGAATTCOCG	3156
Db	3222	CCCGAGCTTGGCAGGCGAGGAGCCAGAAAGCTGTGGGGAACCCCTTGTGTGAATTCOCG	3281
QY	3157	CCGACCCCTTCCCGGTGCGCCCCCGGTGAGGTATCAAAAGGCTCCCGGATGATGCGCGGAGCC	3216
Db	3282	CCGACCCCTTCCCGGTGCGCCCCCGGTGAGGTATCAAAAGGCTCCCGGATGATGCGCGGAGCC	3341
QY	3217	CTCAGCCTTCTCTCACTCACTCACTGTGTCAACCACTGCTCCCTGTGGGCTTCATGACATGCG	3276
Db	3342	CTCAGCCTTCTCTCACTCACTCACTGTGTCAACCACTGCTCCCTGTGGGCTTCATGACATGCG	3401
QY	3277	CCGGGCGGTCCTGCGCGGCCAACCCACATCTCAACCGGCGCTTCCCTCATTTCTCTTGC	3336

Db 3402 CCGGCCGTCCTGCGCGCGCCCAACCATCTTCACACCGGCTCCCTCATCTCTCTGC 3461  
Qy 3337 CAAGCACCACAGCGCTCTTGAGAGGCAATAAGTGCCATCTCCAAAGAAATGTGGTCCA 3396  
Db 3462 CAAGCACCACAGCGCTCTTGAGAGGCAATAAGTGCCATCTCCAAAGAAATGTGGTCCA 3521  
Qy 3397 GCTCCACGTCCCGTACTCAGAGCATGCGAAGGCCCGCGTGGCCCTGTCACATGTGGGCT 3456  
Db 3522 GCTCCACGTCCCGTACTCAGAGCATGCGAAGGCCCGCGTGGCCCTGTCACATGTGGGCT 3581  
Qy 3457 GCCCTGCGCATGAGCCCCAAAAGGTGGCACCTTTAGCGGAGTGAAGCAGAGCAGCT 3516  
Db 3582 GCCCTGCGCATGAGCCCCAAAAGGTGGCACCTTTAGCGGAGTGAAGCAGAGCAGCT 3641  
Qy 3517 GTCCCAAGGGGGCAAGCTGGGCAACCGAGAGCCTGGGGGTGCCCAAGCCCAAGAGGC 3576  
Db 3642 GTCCCAAGGGGGCAAGCTGGGCAACCGAGAGCCTGGGGGTGCCCAAGCCCAAGAGGC 3701  
Qy 3577 GTCCGTGCTGAGAGGGAAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCAAAAGGCAT 3636  
Db 3702 GTCCGTGCTGAGAGGGAAGCTCTGGGCTCAGTTCCGGGCGGAAGCATCAAAAGGCAT 3761  
Qy 3637 TCCCAAGCACAAGGGTGCCTCGGACAGGCGCATCAATACGCGGCTCATCAACCGG 3696  
Db 3762 TCCCAAGCACAAGGGTGCCTCGGACAGGCGCATCAATACGCGGCTCATCAACCGG 3821  
Qy 3697 CACGCAAGCTGACGTCTGTATCAAGGGCACCATCAACAGATCATCGGAGAGACGCC 3756  
Db 3822 CACGCAAGCTGACGTCTGTATCAAGGGCACCATCAACAGATCATCGGAGAGACGCC 3881  
Qy 3757 GAGTGGCTTTGACCCCGGCGGGGAGAGCAGCTGCCCCAAAGGGCACATCTTACGAAGG 3816  
Db 3882 GAGTGGCTTTGACCCCGGCGGGGAGAGCAGCTGCCCCAAAGGGCACATCTTACGAAGG 3941  
Qy 3817 CAAGAAAGGGCCAGCTTGTCTTATGAGGGTGAATGTGTGACCGAGTGTCCAAAGG 3876  
Db 3942 CAAGAAAGGGCCAGCTTGTCTTATGAGGGTGAATGTGTGACCGAGTGTCCAAAGG 4001  
Qy 3877 GAGCAGGAGAAAGCAGCTCAGACCCGCCATGAGACGGCCGCCCAAGCGACCTATGA 3936  
Db 4002 GAGCAGGAGAAAGCAGCTCAGACCCGCCATGAGACGGCCGCCCAAGCGACCTATGA 4061  
Qy 3937 CATGATGAGAGGGCGCGTGGGAGAGACCATCTCTCAAGCAGCATGAAAGTCTCATGG 3996  
Db 4062 CATGATGAGAGGGCGCGTGGGAGAGACCATCTCTCAAGCAGCATGAAAGTCTCATGG 4121  
Qy 3997 CGGTCATCTCCGCGGAGCGACAGAGCCGCCACCATCAAAAGAGCAGACCATCG 4056  
Db 4122 CGGTCATCTCCGCGGAGCGACAGAGCCGCCACCATCAAAAGAGCAGACCATCG 4181  
Qy 4057 CGGTCATCTCAACAAGGGATCCCTCGTCTTACGTGAGGAGCAGAGAGCTACCTGG 4116  
Db 4182 CGGTCATCTCAACAAGGGATCCCTCGTCTTACGTGAGGAGCAGAGAGCTACCTGG 4241  
Qy 4117 TCGGAGAGCCAAAGCTCTTAAAGCGGAGAGGACGCTTCGCCCCCAAGCGCTTCAAGG 4176  
Db 4242 TCGGAGAGCCAAAGCTCTTAAAGCGGAGAGGACGCTTCGCCCCCAAGCGCTTCAAGG 4301  
Qy 4177 CTTGACCGAGGCTTCAAGACGACGAGCCCTGGGCCCCCTGAAGCTGAAGCGGGCATGA 4236  
Db 4302 CTTGACCGAGGCTTCAAGACGAGCCCTGGGCCCCCTGAAGCTGAAGCGGGCATGA 4361  
Qy 4237 GGGCTGTGTGGCAAGTGAAGAGGCGGGCGCTCATTCATGAGATTCGCGGAGGA 4296  
Db 4362 GGGCTGTGTGGCAAGTGAAGAGGCGGGCGCTCATTCATGAGATTCGCGGAGGA 4421  
Qy 4297 GCTGGGGCAGAGCGCCGAGCTGCCCTGGCCCCCGGCGCTCAAGAGAGGCTTCATCAC 4356  
Db 4422 GCTGGGGCAGAGCGCCGAGCTGCCCTGGCCCCCGGCGCTCAAGAGAGGCTTCATCAC 4481  
Qy 4357 GCAGGGCAGCCCGCTCAAGTACAGACCGGGCGCTCAACTGTGCTCAAAAACAGAGA 4416  
Db 4482 GCAGGGCAGCCCGCTCAAGTACAGACCGGGCGCTCAACTGTGCTCAAAAACAGAGA 4541

Qy 4417 GGTACGCTTCCTCATGCGCAGCGCCCGGAGCGTTCCACCCGTGCAACCGCTGATGT 4476  
Db 4542 GGTACGCTTCCTCATGCGCAGCGCCCGGAGCGTTCCACCCGTGCAACCGCTGATGT 4601  
Qy 4477 GATGGCCGACCGCGGGGAGCTGGAAAGCGGCTGCTCAAGAGAGGCTGAAGAGCGGGC 4536  
Db 4602 GATGGCCGACCGCGGGGAGCTGGAAAGCGGCTGCTCAAGAGAGGCTGAAGAGCGGGC 4661  
Qy 4537 AGGACCGCCAGAGCTCGGGGGGCTCATTTGCGGCGGAGCGCCCGGTCAATGTGCTGA 4596  
Db 4662 AGGACCGCCAGAGCTCGGGGGGCTCATTTGCGGCGGAGCGCCCGGTCAATGTGCTGA 4721  
Qy 4597 GCTGGTAAAGCTCGCGGAGAGCGCCCTTGAAGCTATGAGAGCAACGGGGCAACCTTTGCCG 4656  
Db 4722 GCTGGTAAAGCTCGCGGAGAGCGCCCTTGAAGCTATGAGAGCAACGGGGCAACCTTTGCCG 4781  
Qy 4657 CCACCTCCCAAGAGTTGGCCCTGACCATGCGGAGGCCACGCGGCGCTGAGAGAGG 4716  
Db 4782 CCACCTCCCAAGAGTTGGCCCTGACCATGCGGAGGCCACGCGGCGCTGAGAGAGG 4841  
Qy 4717 CAGCTTTGCTCAGCAAGCATCCAGAGCCGAAGCTGACGTGACGCTCGTGAAT 4776  
Db 4842 CAGCTTTGCTCAGCAAGCATCCAGAGCCGAAGCTGACGTGACGCTCGTGAAT 4901  
Qy 4777 CGCCAAAGTCCCGCAGACACCGTGCAGAGACCAACCAACCCATCTGCGCTATGA 4836  
Db 4902 CGCCAAAGTCCCGCAGACACCGTGCAGAGACCAACCAACCCATCTGCGCTATGA 4961  
Qy 4837 GCACCTGTTTGGGGCGTGAAGTGGGAGGAGCTGTATGAGAGCCACATCCCGTGGCTT 4896  
Db 4962 GCACCTGTTTGGGGCGTGAAGTGGGAGGAGCTGTATGAGAGCCACATCCCGTGGCTT 5021  
Qy 4897 GCACCCCACTTCATACCCCGGAGCATCTGTGAGCAGACCGGTGCTTACTACTGCC 4956  
Db 5022 GCACCCCACTTCATACCCCGGAGCATCTGTGAGCAGACCGGTGCTTACTACTGCC 5081  
Qy 4957 CGACACCTTGCGCCCAACCCCAACCTTACCGGACCTGTATCCCACTTACTGCGGG 5016  
Db 5082 CGACACCTTGCGCCCAACCCCAACCTTACCGGACCTGTATCCCACTTACTGCGGG 5141  
Qy 5017 CTACCCCGACAGCGGGGCGTGAAGAACGGGACGACATCAATGATGATCATCACTTC 5076  
Db 5142 CTACCCCGACAGCGGGGCGTGAAGAACGGGACGACATCAATGATGATCATCACTTC 5201  
Qy 5077 GCAGCAGATGACCAACAACAGGCCACCGCATGGCCAGAGAGCTGATATGAGGGG 5136  
Db 5202 GCAGCAGATGACCAACAACAGGCCACCGCATGGCCAGAGAGCTGATATGAGGGG 5261  
Qy 5137 CTTCTGCGCCGCGAGTCTGCTGGCACTCACTACGCTGCGGATCCCGAGGATCAT 5196  
Db 5262 CTTCTGCGCCGCGAGTCTGCTGGCACTCACTACGCTGCGGATCCCGAGGATCAT 5321  
Qy 5197 GCACTGTGCCAAGTGGCACAACGTGCTGTGTGTCGTGCCCCGACACAGGACCCGAGC 5256  
Db 5322 GCACTGTGCCAAGTGGCACAACGTGCTGTGTGTCGTGCCCCGACACAGGACCCGAGC 5381  
Qy 5257 CACGCGCATGAGACGCGCTTGGCTTACCTCCACAGCGCGCCAGCGCTTCAAGCGCGCA 5316  
Db 5382 CACGCGCATGAGACGCGCTTGGCTTACCTCCACAGCGCGCCAGCGCTTCAAGCGCGCA 5441  
Qy 5317 CAGCAGCTTCCCACTTCTCCAGAGAGTCAACAACCTTGAACAAAACAACAACAGCTC 5376  
Db 5442 CAGCAGCTTCCCACTTCTCCAGAGAGTCAACAACCTTGAACAAAACAACAACAGCTC 5501  
Qy 5377 CTGCTGCGAGGGGAGCGAGACCGGATCGAGAGCGGGAACCGGGATTCGGAGCGGGAAA 5436  
Db 5502 CTGCTGCGAGGGGAGCGAGACCGGATCGAGAGCGGGAACCGGGATTCGGAGCGGGAAA 5561  
Qy 5437 GTCCATCTCTCAAGTCAACAAGACGATGAGACAGCAACCATCTTGAAGACCTGTATCAGA 5496  
Db 5562 GTCCATCTCTCAAGTCAACAAGACGATGAGACAGCAACCATCTTGAAGACCTGTATCAGA 5621



QY	5497	CGAAGCAGCGGCGAGCAAGCGGACAGAGGGGGGGTGGGGGGCGAGCAAGCGCGCCCGC	5556
Db	5622	GCAGAGCAGCGCGAGCAAGCGGACAGAGGGGGGGTGGGGGGCGAGCAAGCGCGCGCCCGC	5661
QY	5557	CTCCCACTCCCATGGCCCAACAGCACTCGCCCATCTCCCTCGGAGCCAGATGCGCTTCA	5616
Db	5662	CTCCCACTCCCATGGCCCAACAGCACTCGCCCATCTCCCTCGGAGCCAGATGCGCTTCA	5741
QY	5617	GCAGAGCCCAAGTGTGCTTCACAACAGGCAATGAGGGTATCATCACCGCTGTGAGCC	5676
Db	5742	GCAGAGCCCAAGTGTGCTTCACAACAGGCAATGAGGGTATCATCACCGCTGTGAGCC	5801
QY	5677	CAGCAAGCCCAAGGTCTGAGGTCCACTCACTCTCTCAACCCGTTGGCCCAAGTGTGAC	5736
Db	5802	CAGCAAGCCCAAGGTCTGAGGTCCACTCACTCTCTCAACCCGTTGGCCCAAGTGTGAC	5861
QY	5737	ATTCCCACTTGCACCCCACTGTGCCACTGTGGGGGCAACCTCGATGGAGGTCTACCCCTA	5796
Db	5862	ATTCCCACTTGCACCCCACTGTGCCACTGTGGGGGCAACCTCGATGGAGGTCTACCCCTA	5921
QY	5797	CATGAGCCCGCTCTTTCCTGCCAAGAGAGGCCCGCGGGTTCGCGCGGCGAGAGCGGCGCG	5856
Db	5922	CATGAGCCCGCTCTTTCCTGCCAAGAGAGGCCCGCGGGTTCGCGCGGCGAGAGCGGCGCG	5981
QY	5857	AGCAGACACCGGCGCATGCTTCTCTGCGCAAGCCCGCCAGCCCGCTTCGGAGTGCACCC	5916
Db	5982	AGCAGACACCGGCGCATGCTTCTCTGCGCAAGCCCGCCAGCCCGCTTCGGAGTGCACCC	6041
QY	5917	CTCCCTCCCGCCAGCAAGGGGTCTGAGAGCCCGGGCCCTATGTGCTCTGTCTGTGACAGC	5976
Db	6042	CTCCCTCCCGCCAGCAAGGGGTCTGAGAGCCCGGGCCCTATGTGCTCTGTCTGTGACAGC	6101
QY	5977	CACCATCGCCCGCACCCCTGTGGAAGAACTTGCACCTTCAACGACGCAAGCCCGGACCGCGC	6036
Db	6102	CACCATCGCCCGCACCCCTGTGGAAGAACTTGCACCTTCAACGACGCAAGCCCGGACCGCGC	6161
QY	6037	GGCGGCACCTTCCTTCGCGCTTGGACCCGCGACCGGGAAAAAGCTAAGTAAACCTTTTC	6096
Db	6162	GGCGGCACCTTCCTTCGCGCTTGGACCCGCGACCGGGAAAAAGCTAAGTAAACCTTTTC	6221
QY	6097	CATCAGGAACATTGGAACCTCCGTTCTCTGGGTTTACACCGGAGAGCTACAGCCCGGAAG	6156
Db	6222	CATCAGGAACATTGGAACCTCCGTTCTCTGGGTTTACACCGGAGAGCTACAGCCCGGAAG	6281
QY	6157	GGTGAAGCCCGTCAGCCCTGTGAGCTCAACCAAGTGTGACCCACGACAAAGGGGCTCCCAA	6216
Db	6282	GGTGAAGCCCGTCAGCCCTGTGAGCTCAACCAAGTGTGACCCACGACAAAGGGGCTCCCAA	6341
QY	6217	GCACCTGGAAGAGTCTGACAAAGAGCCACTTGAAGGGAGAGTGTGCGGCCAAGACGACAG	6276
Db	6342	GCACCTGGAAGAGTCTGACAAAGAGCCACTTGAAGGGAGAGTGTGCGGCCAAGACGACAG	6401
QY	6277	CCCGGTGAAGTTTGGCGGGAGGGCGCCCACTTCCACACTTGGGGCGCTGCTGAGAG	6336
Db	6402	CCCGGTGAAGTTTGGCGGGAGGGCGCCCACTTCCACACTTGGGGCGCTGCTGAGAG	6461
QY	6337	CCAGGCCCTGTCAGGCCCGCTGCTCAAGACCGCCCGAGGGGTCAAAGGTCAACGACGGGT	6396
Db	6462	CCAGGCCCTGTCAGGCCCGCTGCTCAAGACCGCCCGAGGGGTCAAAGGTCAACGACGGGT	6521
QY	6397	GGTCAACCTTGGCCGACCATCAAGTGAAGTCAATCAACAGGACTTACACCCCGGACCAACC	6456
Db	6522	GGTCAACCTTGGCCGACCATCAAGTGAAGTCAATCAACAGGACTTACACCCCGGACCAACC	6581
QY	6457	ACAGCAGCTCAGCGCACCCCTTGCAGCGCCCGCTTCTTCTTCCCTTGGGGCAAGTGTGCC	6516
Db	6582	ACAGCAGCTCAGCGCACCCCTTGCAGCGCCCGCTTCTTCTTCCCTTGGGGCAAGTGTGCC	6641
QY	6517	CGTCTGTGACCTTCGCGCGGCGCCACCAAGTGAACCTTCACTCCCGGCCCGGACATGTTGTC	6576
Db	6642	CGTCTGTGACCTTCGCGCGGCGCCACCAAGTGAACCTTCACTCCCGGCCCGGACATGTTGTC	6701
QY	6577	CCCGGCGCGGTCTTCCCGCCACAGCGAAGGGGGCAAGAGTCTTCAAGGCCAAACAGAC	6636

Dp	6702	CCGGCCCCGTGGCTCCCCCCACAGCCGAAGGGGGCGAAGGCTCTCCAGAGCCAAACAAGAC	6761
Qy	6637	GTCCGCTTTGGGTGGTGGTGAAGACGGATTATGAACCTGTGTCCCCACCGAGGGCATGAC	6696
Dp	6762	GTCCGTCTTGGGTGGTGGTGAAGACGGATTATGAACCTGTGTCCCCACCGAGGGCATGAC	6821
Qy	6697	GGAGCCAGGGGCACTCCCGGAGTGTGTGTACCCGCTCTGTACCGGGGATGGGGACAAGAC	6756
Dp	6822	GGAGCCAGGGGCACTCCCGGAGTGTGTGTACCCGCTCTGTACCGGGGATGGGGACAAGAC	6881
Qy	6757	GGAGCCAGGACAGATGGGGCTCAAGTCTCCAGGGCAACACGAGCCAGCGCCAGCTCTT	6816
Dp	6882	GGAGCCAGGACAGATGGGGCTCAAGTCTCCAGGGCAACACGAGCCAGCGCCAGCTCTT	6941
Qy	6817	CAGCAAGCTGACCGAGAGCACTCCGCAATGCTCAAGTCCAGAGAGCAAGATCAACA	6876
Dp	6942	CAGCAAGCTGACCGAGAGCACTCCGCAATGCTCAAGTCCAGAGAGCAAGATCAACA	7001
Qy	6877	GAGCTGAACAACCCACAACCGGAATGAGCTGAATACATATACAGCCGCTGGGACGGA	6936
Dp	7002	GAGCTGAACAACCCACAACCGGAATGAGCTGAATACATATACAGCCGCTGGGACGGA	7061
Qy	6937	GATTTTCAATATGCCCGCCCAATCCCGGAACAAGGCTTATACCTATGAAGGACGGGGGT	6996
Dp	7062	GATTTTCAATATATGCCCGCCCAATCCCGGAACAAGGCTTATGAAGGACGGGGGT	7121
Qy	6997	GCAAGGAACATGCGCAGCAACAACATGGGGCTGAGAGCCATTAATTGAAGGACCTCATGAG	7056
Dp	7122	GCAAGGAACATGCGCAGCAACAACATGGGGCTGAGAGCCATTAATTGAAGGACCTCATGAG	7181
Qy	7057	TAAATATGACCAAGTGGGAGAAGTCCCGCGCTCAGCGCCAAATGCTTTTAAACCTCTGAA	7116
Dp	7182	TAAATATGACCAAGTGGGAGAAGTCCCGCGCTCAGCGCCAAATGCTTTTAAACCTCTGAA	7241
Qy	7117	TGCCAATGCGCAGCTGCGCCCTGTGTATGCCCATTAACCGCTGTGACGGAAGGATGACCA	7176
Dp	7242	TGCCAATGCGCAGCTGCGCCCTGTGTATGCCCATTAACCGCTGTGACGGAAGGATGACCA	7301
Qy	7177	CACACTCACTCGCGCAGGTGGCGGGGGGGAAGGCCAAGGTCTCTGGCAGAACCCAGACACCG	7236
Dp	7302	CACACTCACTCGCGCAGGTGGCGGGGGGGAAGGCCAAGGTCTCTGGCAGAACCCAGACACCG	7361
Qy	7237	AAAAGCCAAGTCCCCCGGCCCCCGGCGCTGAGCATTGGGGACCGGCAACCTCTGTCTCTC	7296
Dp	7362	AAAAGCCAAGTCCCCCGGCCCCCGGCGCTGAGCATTGGGGACCGGCAACCTCTGTCTCTC	7421
Qy	7297	AGTGCATCTCGGAGGGAGACTGCAACCGCGCGGACGCGCTCAACCAACCGCGCTGTGGAGGA	7356
Dp	7422	AGTGCATCTCGGAGGGAGACTGCAACCGCGCGGACGCGCTCAACCAACCGCGCTGTGGAGGA	7481
Qy	7357	CAGGCCCTCGTCCGCGAGGTTCCAGGCCATTCCCTTAACACCCCTGTATCATATGCGGCTGCA	7416
Dp	7482	CAGGCCCTCGTCCGCGAGGTTCCAGGCCATTCCCTTAACACCCCTGTATCATATGCGGCTGCA	7541
Qy	7417	GGCGGGGTGATGAGCTTCCCCCAACCCCAACCGGCGCTTCCCGCGGGCAAGCGGCGCTCGC	7476
Dp	7542	GGCGGGGTGATGAGCTTCCCCCAACCCCAACCGGCGCTTCCCGCGGGCAAGCGGCGCTCGC	7601
Qy	7477	TGGCCCCCACAAGCTGTGGACCGAGAGGCCAAGCCACTGTCTGTGCTGCGATGACGAGAC	7536
Dp	7602	TGGCCCCCACAAGCTGTGGACCGAGAGGCCAAGCCACTGTCTGTGCTGCGATGACGAGAC	7661
Qy	7537	ACTCTCCGACAGGAGTGACTCAAGAAACAGGGCGGGGGGGGGGGGGGGGGGTGUBSPRLUG	7596
Dp	7662	ACTCTCCGACAGGAGTGACTCAAGAAACAGGGCGGGGGGGGGGGGGGGGGGTGUBSPRLUG	7711
Qy	7597	TCAAGTCCCAAGCAGCCACAAGAACGGCCCTGTGACGAGACCGGCGGCTGCGACTCCCCC	7656
Dp	7712	TCAAGTCCCAAGCAGCCACAAGAACGGCCCTGTGACGAGACCGGCGGCTGCGACTCCCCC	7771
Qy	7657	AACCAAGAAAGAGCCCTGAGTCCGCTGCGCTCATTCATCTGTCCGTCCAGAGCCG	7716



[illegible]

Db	6639	TGGTGGGCTCCAGACTGTGTCCGAGGTGAAGACTTCTTACTTCACTTACAGAAAGGCG	6580
Qy	1980	AGAACTTGATGAGATCTTSCAGACGACAGCTGAAGATGAGAGAGAGAGAACGCGC	2039
Db	6579	AGAACTTGATGAGATCTTSCAGACGACAGCTGAAGATGAGAGAGAGAGAACGCGC	6520
Qy	2040	GGAGGAAAGAGAGAAAGCGCCCGCGCGCGCCAGCCGAGAGAGGCTGCATTCCCGCGCGTGG	2059
Db	6519	GGAGGAAAGAGAGAAAGCGCCCGCGCGCGCCAGCCGAGAGAGGCTGCATTCCCGCGCGTGG	6460
Qy	2100	TGAGGATGAGAGATGAGAGCGCTCGGCGCTGAGCGGAAATGAGAGAGAGATGCTGAGG	2159
Db	6459	TGAGGATGAGAGATGAGAGCGCTCGGCGCTGAGCGGAAATGAGAGAGAGATGCTGAGG	6400
Qy	2160	AGGCTGAAGCCTTACATGCTCTTGAGGAATGAGGTGCCAGAGGGGAAATGCAATGGCCAG	2219
Db	6399	AGGCTGA-----AG	6391
Qy	2220	CCACTGTCAACACAGCTTCAGACACCGAGAGCATCCCTCTCTCTCAACATGAGGCGGCA	2279
Db	6390	CCACTGTCAACACAGCTTCAGACACCGAGAGCATCCCTCTCTCTCAACATGAGGCGGCA	6331
Qy	2280	AGGACACAGGGGAGAAATGGGCGCCAGAGCCCGCAGCACCTTGAGGGGCGGACGAGGCGACCC	2339
Db	6330	AGGACACAGGGGAGAAATGGGCGCCAGAGCCCGCAGCACCTTGAGGGGCGGACGAGGCGACCC	6271
Qy	2340	CAGGCGCACCCACCCACACACGAGAGCATCCCGGGCGCCCATTTGAGCCACCCCGGCT	2399
Db	6270	CAGGCGCACCCACCCACACACGAGAGCATCCCGGGCGCCCATTTGAGCTCACCCCGGCT	6211
Qy	2400	CTGAAGCCACCGGAGCCCTTACGCGCCCAACAGACATCCCATGCGCTCTGCACTTCTC	2459
Db	6210	CTGAAGCCACCTTATAGCCCTTACGCGCCCAACAGACATCCCATTTCTCTTCACTTCTC	6151
Qy	2460	CTGTGTCTCCCAAGAGAGAGAGAGAGAGACCGGAGAGCGCCCACTGAGGAGG	2519
Db	6150	CTGTGTCTCCCAAGAGAGAGAGAGAGAGACCGGAGAGCGCCCACTGAGGAGG	6091
Qy	2520	GGGAGAGAGAGAAAGCCCGCGGCTTGAGAGCTGAGCATGTGACACAGAGGAGGCGAGG	2579
Db	6090	GGGAGAGAGAGAAAGCCCGCGGCTTGAGAGCTGAGCATGTGACACAGAGGAGGCGAGG	6031
Qy	2580	AGCCCTGACAGAGGAGTGCACGAGAGAGCCGAGAGGAGGCGGCGCAAGGGCACAGACG	2639
Db	6030	AGCCCTGACAGAGGAGTGCACGAGAGAGCCGAGAGGAGGCGGCGCAAGGGCACAGACG	5971
Qy	2640	CGAGAGCGGCTTGAGAGCCACGCGCGAGGGGCGCTCAAGGACAGAGAAAGAGAGGCGGGA	2699
Db	5970	CGAGAGCGGCTTGAGAGCCACGCGCGAGGGGCGCTCAAGGACAGAGAAAGAGAGGCGGGA	5911
Qy	2700	GGGACAGAGGACACCATGCGCAAGAGCTCGGAGCGCGCCCGACAGACAGGCACTCCAGTCTA	2759
Db	5510	GGGACAGAGGACACCATGCGCAAGAGCTCGGAGCGCGCCCGACAGACAGGCACTCCAGTCTCA	5851
Qy	2760	CTTGCACTGCAGAGAGAGTGAAGGCGAGAGGCGGCGACAAAGACCGGCTGTCTCTCC	2819
Db	5850	CTTGCACTGCAGAGAGAGTGAAGGCGAGAGGCGGCGACAAAGAACCGGCTGTCTCTCC	5791
Qy	2820	CAAGGCCACGCTTCTTCACTCCCACTGAGCGACCCCGGGGCAATGCTTCACTCCCAAGAC	2879
Db	5790	CAAGGCCACGCTTCTTCACTCCCACTGAGCGACCCCGGGGCAATGCTTCACTCCCAAGAC	5731
Qy	2880	CACGTGAACCTGAAGCAGCTGAAGCAGGAGGAGGCTGCATTCGCCCATCAGGCTTCACCA	2939
Db	5730	CACGTGAACCTGAAGCAGCTGAAGCAGGAGGAGGCTGCATTCGCCCATCAGGCTTCACCA	5671
Qy	2940	AACTTCATGAGCCCCCGGAGAGGACGACGCTTCCACCAAGCAGCTTCCCGACGCCAC	2999
Db	5670	AACTTCATGAGCCCCCGGAGAGGACGACGCTTCCACCAAGCAGCTTCCCGACGCCAC	5611
Qy	3000	CGCCACCGCAAAACCTGACACCCGAGAGCGACCGCCCTCAGCAGCCTTGGGACGAGGCCCG	3059

Db	5610	CGCCACCGCAAAACCTGACGCGGAGACGACGCCCCCTCAGAGCTTGGCAGACGCCCC	5551
Oy	3060	GGGGCAAGACAGAGAGCCCGGACACCCCGCGCAAGAGAG-----	3100
Db	5550	GGGGCAAGACAGAGAGCCCGGACACCCCGCGCAAGAGAGAGCCTGTCTT	5491
Oy	3101	-----GCCCTTGGCAGCGGAGGCGCCAGAAAGCTGCGCTGGGAGACCCCTTGTGTGAATTCCG	3155
Db	5490	TCCAGCCTTGGCAGCGGAGGCGCCAGAAAGCTGCTGGGAGACCCCTTGTGTGAATTCCG	5431
Oy	3156	GCTTGCCTTCCCGGTGCCCCCGCTGAGGTGATCAAGGCTCCCGCATGCCCGAGC	3215
Db	5430	GCTTGCCTTCCCGGTGCCCCCGCTGAGGTGATCAAGGCTCCCGCATGCCCGAGC	5371
Oy	3216	CCTAGCCTTCTCTTAGAGCTTCACTGTGATCCACTGTGCCCCCTGGGCTCTCATGACATG	3275
Db	5370	CCTAGCCTTCTCTTAGAGCTTCACTGTGATCCACTGTGCCCCCTGGGCTCTCATGACATG	5311
Oy	3276	CCCCCGCCGTCTGTGCGGCGCCACCCACATCTCCAAACCGGCTCCCTCATCTCTTG	3335
Db	5310	CCCCCGCCGTCTGTGCGGCGCCACCCACATCTCCAAACCGGCTCCCTCATCTCTTG	5251
Oy	3336	CCAAGCACCCGAGCGTCTCCGAGAGAGGAAATAGGTGCATCTCCCAAGAAATGTGCTGC	3395
Db	5250	CCAAGCACCCGAGCGTCTCCGAGAGGAAATAGGTGCATCTCCCAAGAAATGTGCTGC	5191
Oy	3396	AGCTTCAGCTCCCGTACTCAGAGCATGCGAAAGGCCCGGTGGGCGCTGTCACTATGGGCG	3455
Db	5190	AGCTTCAGCTCCCGTACTCAGAGCATGCGAAAGGCCCGGTGGGCGCTGTCACTATGGGCG	5131
Oy	3456	TGCCCCCTTGCCTATGGAACCCCAAAAAGCTGTGCACCTTCAAGGAGTGAAGCAGAGCACG	3515
Db	5130	TGCCCCCTTGCCTATGGAACCCCAAAAAGCTGTGCACCTTCAAGGAGTGAAGCAGAGCACG	5071
Oy	3516	TGTCCCAACGGGGCCAGGCTGGGCGCACCGAGAGCTTGGGGGTGCCACAGCCAGAGG	3575
Db	5070	TGTCCCAACGGGGCCAGGCTGGGCGCACCGAGAGCTTGGGGGTGCCACAGCCAGAGG	5011
Oy	3576	CGTCCGTCTGAGAGAGGACAAGCTTGTGGGCTCAGTTCCGGGCGGAAAGCATCAACAAAGCA	3635
Db	5010	CGTCCGTCTGAGAGAGGACAAGCTTGTGGGCTCAGTTCCGGGCGGAAAGCATCAACAAAGCA	4951
Oy	3636	TTCCCGACACAGGGTGCCCTGTGGACAGCGCATATACATACCGGGCTTCATCACCCAG	3695
Db	4950	TTCCCGACACAGGGTGCCCTGTGGACAGCGCATATACATACCGGGCTTCATCACCCAG	4891
Oy	3696	GACGCGCAGCTGACGTCCTCTGTAAAGGGCACCATACCAAGATATATGGCGAGAGACGCC	3755
Db	4890	GACGCGCAGCTGACGTCCTCTGTAAAGGGCACCATACCAAGATATATGGCGAGAGACGCC	4831
Oy	3756	CGAGTCGCTTGGACCGCGCGCGGGAGAGCAGCTTGCACAGGGCCACCTCATCTTACAGAG	3815
Db	4830	CGAGTCGCTTGGACCGCGCGCGGGAGAGCAGCTTGCACAGGGCCACCTCATCTTACAGAG	4771
Oy	3816	GCAGAGAGGGCCACGTCCTTGTCTATGAGGGTGCGATGTCTGTGACCCAGTGTCCAAAG	3875
Db	4770	GCAGAGAGGGCCACGTCCTTGTCTATGAGGGTGCGATGTCTGTGACCCAGTGTCCAAAG	4711
Oy	3876	AGGACCGGAGAAAGCAGCTCAGAGACCCCCCCCATATGAGACGGCGCGCCCGCCAGGACCTATG	3935
Db	4710	AGGACCGGAGAAAGCAGCTCAGAGACCCCCCCCATATGAGACGGCGCGCCCGCCAGGACCTATG	4651
Oy	3936	ACATGATGAGAGGGCGCGGTGGGACAGAGCATCTCTCAGCGACGATCGAAGGTCTACATGG	3995
Db	4650	ACATGATGAGAGGGCGCGGTGGGACAGAGCATCTCTCAGCGACGATCGAAGGTCTACATGG	4591
Oy	3996	GCCGTGCCATCCCGCGGAGCGACACAGCCCCACACCTTCAAAGAGCAGACCAATCC	4055
Db	4590	GCCGTGCCATCCCGCGGAGCGACACAGCCCCACACCTTCAAAGAGCAGACCAATCC	4531
Oy	4056	GCGGTCATCAACACAGAGATCCCTCGGTCTTACTGTGAGGACACAGAGACTTACTGTGC	4115
Db	4530	GCGGTCATCAACACAGAGATCCCTCGGTCTTACTGTGAGGACACAGAGACTTACTGTGC	4471

QY	4116	GTCCGAGAGCCAACTCTCTAAAGCGGAGGCAAGCTCCGCGCCCAACCGCTCTACAGG	4175
Db	4470	GTCGGAGAGCCAACTCTCTAAAGGGAGGCAAGCTCCGCGCCCAACCGCTCTACAGG	4411
QY	4176	ACCTGACCGAAGCCTTACAAGAAGCAGAGGCCCTCGGGCCCCCTGAAGCTGAAGCGGGCCATG	4235
Db	4410	ACCTGACCGAAGCCTTACAAGAAGCAGAGGCCCTCGGGCCCCCTGAAGCTGAAGCGGGCCATG	4351
QY	4236	AGGACCTGTGTGGCCACGAGTGAAGAGGCGGGCGCTTCATTCATGAGATCCCGCGAGG	4295
Db	4350	AGGACCTGTGTGGCCACGAGTGAAGAGGCGGGCGCTTCATTCATGAGATCCCGCGAGG	4291
QY	4296	AGCTGCGGCAACAGCCCCGAGGTGCCCCCTGGCCCCGGGGCGGCTCAAGAAAGGAGCTCCATCA	4355
Db	4290	AGCTGCGGCAACAGCCCCGAGGTGCCCCCTGGCCCCGGGGCGGCTCAAGAAAGGAGCTCCATCA	4231
QY	4356	CGCAGAGGACCCCGCTCAAGTACGACACGCGCGCTGCACACCTAGTGCCTCCAAAAGCAGC	4415
Db	4230	CGCAGAGGACCCCGCTCAAGTACGACACGCGCGCTGCACACCTAGTGCCTCCAAAAGCAGC	4171
QY	4416	ACGTACGCTCCCTCATGTGGCAGAGCCCCGGGCGGAGAGTTTCCACCCGTGTACACCCGTGGATG	4475
Db	4170	ACGTACGCTCCCTCATGTGGCAGAGCCCCGGGCGGAGAGTTTCCACCCGTGTACACCCGTGGATG	4111
QY	4476	TGATGCGCGAAGCCCCGGGCACTGGAAAGTGCCTGTCTACAGAGAGAGCCTGAAGAGCGGAG	4535
Db	4110	TGATGCGCGAAGCCCCGGGCACTGGAAAGTGCCTGTCTACAGAGAGAGCCTGAAGAGCGGAG	4051
QY	4536	CAGGAGCCGCGACGAGCTCGGGGGGCTTCATTGCGCGGGCGGCCCGGTCAATGTGCTTG	4595
Db	4050	CAGGAGCCGCGACGAGCTCGGGGGGCTTCATTGCGCGGGCGGCCCGGTCAATGTGCTTG	3991
QY	4596	AGCTGGGTAAAGCGCGGCGAGAGCCCCCTGACCTATAGAGACCAAGGGGACCCCTTTCGCG	4655
Db	3990	AGCTGGGTAAAGCGCGGCGAGAGCCCCCTGACCTATAGAGAGCAAGGGGACCCCTTTCGCG	3931
QY	4656	GCCAACCTTCCACGAGGTTGCGCCGTGACCAATGCGGAGCCCAAGCGCGGCTCGAGAGG	4715
Db	3930	GCCAACCTTCCACGAGGTTGCGCCGTGACCAATGCGGAGCCCAAGCGCGGCTCGAGAGG	3871
QY	4716	GCAACCTTTGCTCCAGCAAGGCAATCCCAAGACCGAAGCTGAAGTGCAGAGCCCTCGTAGA	4775
Db	3870	GCAACCTTTGCTCCAGCAAGGCAATCCCAAGACCGAAGCTGAAGTGCAGAGCCCTCGTAGA	3811
QY	4776	TCGCGAAGTCCCGCGACAGACCGTGCGCCAGACCAACCAACCCCATCTCGCCATATG	4835
Db	3810	TCGCGAAGTCCCGCGACAGACCGTGCGCCAGACCAACCAACCCCATCTCGCCATATG	3751
QY	4836	AGCACTTGCTTCGGGGCGTGAATGAGCGTGAACCTGTATTCGAGCCACATCCCTTGAGCT	4895
Db	3750	AGCACTTGCTTCGGGGCGTGAATGAGCGTGAACCTGTATTCGAGCCACATCCCTTGAGCT	3691
QY	4896	TCGACCCCAACCTCCATACCCCGGGGCAATCCCTGACGAGCGAGCGGCTGACCTACATCGC	4955
Db	3690	TCGACCCCAACCTCCATACCCCGGGGCAATCCCTGACGAGCGAGCGGCTGACCTACATCGC	3631
QY	4956	CCCGACACTTGAGCCCCCAACCCCAACCTTACCGGACCTGTATCCGACCTCACTCATCGCG	5015
Db	3630	CCCGACACTTGAGCCCCCAACCCCAACCTTACCGGACCTGTATCCGACCTCACTCATCGCG	3571
QY	5016	GCTACCCCGACACGGCGCGCGCTGTAGAAACCGGCGAGACCAATCATATGATCTACATCACT	5075
Db	3570	GCTACCCCGACACGGCGCGCGCTGTAGAAACCGGCGAGACCAATCATATGATCTACATCACT	3511
QY	5076	CGCAGCAGATGACCAACAACGGCCACCGCCATAGCCCCAGCAGAGCTGATGCTGAGGG	5135
Db	3510	CGCAGCAGATGACCAACAACGGCCACCGCCATAGCCCCAGCAGAGCTGATGCTGAGGG	3451
QY	5136	GCTCTTCGCCCCGAGGTCTCGCTGGCACTCAATACGCTGCGGGTCCCGAGGACATCA	5195
Db	3450	GCTCTTCGCCCCGAGGTCTCGCTGGCACTCAATACGCTGCGGGTCCCGAGGACATCA	3391

QY	5196	TCGACCTGTCCCAAGTGCACA	CCCTGCTGTGTGTGCCCCCGACACACAGCACCCCAAG	5255				
Dp	3390	TCGACCTGTCCCAAGTGCACA	CCCTGCTGTGTGTGCCCCCGACACACAGCACCCCAAG	3331				
QY	5256	CCACCGCATAGAACCGGCTTGCTAC	TCTGCTGTGCCCCCGAGCCCTTCAGACGCGCC	5315				
Dp	3330	CCACCGCATAGAACCGGCTTGCTAC	TCTGCTGTGCCCCCGAGCCCTTCAGACGCGCC	3271				
QY	5316	ACAGCAGCTCCCACTCTCC	CCAGAGGTCCAAACA	CTTGA	CAAAACA	CAACA	CAACAGT	5375
Dp	3270	ACAGCAGCTCCCACTCTCC	CCAGAGGTCCAAACA	CTTGA	CAAAACA	CAACA	CAACAGT	3211
QY	5376	CCTGTCTCGAGCGGGAGGAGAC	CGGGATTCGAGACGGGAC	CCGGGATCTGGGAGCGGGAAA	5435			
Dp	3210	CCTGTCTCGAGCGGGAGGAGAC	CGGGATTCGAGACGGGAC	CCGGGATCTGGGAGCGGGAAA	3151			
QY	5436	AGTCCATCTCAACGTCACA	CGAGGAGACACGAC	CCCATCTGGAGAC	CTTGTTACAG	5495		
Dp	3150	AGTCCATCTCAACGTCACA	CGAGGAGACACGAC	CCCATCTGGAGAC	CTTGTTACAG	3091		
QY	5496	AGCAGAGCAGCGGAGCAGCGG	CAGCAGCGGCGGGGTGGGGAGCAGCAGCGCCCG	5555				
Dp	3090	AGCAGAGCAGCGGAGCAGCGG	CAGCAGCGGCGGGGTGGGGAGCAGCAGCGCCCG	3031				
QY	5556	CCTCCCACTCCCATGCCCCA	CGACACTCGCCCATCTCCCTCGAGACCCAGGATCCCTCC	5615				
Dp	3030	CCTCCCACTCCCATGCCCCA	CGACACTCGCCCATCTCCCTCGAGACCCAGGATCCCTCC	2971				
QY	5616	AGCAGAGCCCAAGTGTGTTACA	CAACAGGCATGAAAGGATATCA	CAACCGCTGTGAGC	5675			
Dp	2970	AGCAGAGCCCAAGTGTGTTACA	CAACAGGCATGAAAGGATATCA	CAACCGCTGTGAGC	2911			
QY	5676	CCAGCAAGCCCAAGGCTCTG	AGGTCCACTCCATCTCTCA	CCCGGTTGCGCCAC	CTGACA	5735		
Dp	2910	CCAGCAAGCCCAAGGCTCTG	AGGTCCACTCCATCTCTCA	CCCGGTTGCGCCAC	CTGACA	2851		
QY	5736	CATTCCCACTGCGCACCCCA	CTGCGCACCTTGCGGCGGAC	CCCTGATGGGCTCTAC	CCCTACCC	5795		
Dp	2850	CATTCCCACTGCGCACCCCA	CTGCGCACCTTGCGGCGGAC	CCCTGATGGGCTCTAC	CCCTACCC	2791		
QY	5796	TCAATGAGACCCGCTCTTG	GCTGCCCAAGAGAGGCCCCCGGGGTGCCCCCGACAGAGCGGCCCC	5855				
Dp	2790	TCAATGAGACCCGCTCTTG	GCTGCCCAAGAGAGGCCCCCGGGGTGCCCCCGACAGAGCGGCCCC	2731				
QY	5856	GAGCAGACACCGGCGCATG	CTTCCCTCGACAGCCCGGACCCCGGCTCGGGGTGAGACCGG	5915				
Dp	2730	GAGCAGACACCGGCGCATG	CTTCCCTCGACAGCCCGGACCCCGGCTCGGGGTGAGACCGG	2671				
QY	5916	CCTCCTCCCGCAGCAAGG	CTTCGAGGCCCGGCGCTTAGTG	CTCTTGCTCTTG	GGACAG	5975		
Dp	2670	CCTCCTCCCGCAGCAAGG	CTTCGAGGCCCGGCGCTTAGTG	CTCTTGCTCTTG	GGACAG	2611		
QY	5976	CCACCATGCGCCGAC	CCCCCTGCGAAGAACTTCGAC	CTCACACGCGCAC	CGCCGAGCCCGC	6035		
Dp	2610	CCACCATGCGCCGAC	CCCCCTGCGAAGAACTTCGAC	CTCACACGCGCAC	CGCCGAGCCCGC	2551		
QY	6036	CGGCGGCACCTGCGCTCG	CGGCTTCGAGACCCCGGACCGGGAAA	AACATCA	AGTTAA	CCCTTTT	6095	
Dp	2550	CGGCGGCACCTGCGCTCG	CGGCTTCGAGACCCCGGACCGGGAAA	AACATCA	AGTTAA	CCCTTTT	2491	
QY	6096	CCATTCAGAACTCGAACT	CGGTTCTCTGAGTTAC	ACAGCGAGAC	GTTACAG	CGCCCGAAG	6155	
Dp	2490	CCATTCAGAACTCGAACT	CGGTTCTCTGAGTTAC	ACAGCGAGAC	GTTACAG	CGCCCGAAG	2431	
QY	6156	GGGTGAGACCCGTCAG	CCCTGTGAGCTCAC	CCAGTTCGAC	CCACGAC	AGAGGAGCTCCCA	6215	
Dp	2430	GGGTGAGACCCGTCAG	CCCTGTGAGCTCAC	CCAGTTCGAC	CCACGAC	AGAGGAGCTCCCA	2371	
QY	6216	AGCAGCTGGAGAGAGCT	TCGACAAAGGCGCAC	CTGAGAGGGGAGAC	CTMGCGCCAC	AGAGGACG	6275	
Dp	2370	AGCAGCTGGAGAGAGCT	TCGACAAAGGCGCAC	CTGAGAGGGGAGAC	CTMGCGCCAC	AGAGGACG	2311	
QY	6276	GCCCCGTGAACTTGG	CGGGAGGCGCGCAC	CTTCCAC	CACTGCGCGCTG	CTGTGAGA	6335	

Dp	2310	GCCTCGTGAAGCTTTGGCGGGGAGGCCTCCACCTTCCACACTTCGGGCTCGTCTGAGA	2251
Qy	6336	GCCAGCCCTCGTCCAGCCCGCTGCTCCAGACCCGCCAGGGGTCAAAAGTCAACAGCGGG	6395
Dp	2250	GCCAGCCCTCGTCCAGCCCGCTGCTCCAGACCCGCCAGGGGTCAAAAGTCAACAGCGGG	2191
Qy	6396	TGTTGACCCCTGGCCAGACATCATGTAGGTCAATCAACAGAACTAACCTCCGACCAAC	6455
Dp	2190	TGTTGACCCCTGGCCAGACATCATGTAGGTCAATCAACAGAACTAACCTCCGACCAAC	2131
Qy	6456	CACAGCAGCTCAGGGGACCCCTGCGCGCCCTCTACTCTCTTCCCTGGGGCCAGCTGCC	6515
Dp	2130	CACAGCAGCTCAGGGGACCCCTGCGCGCCCTCTACTCTCTTCCCTGGGGCCAGCTGCC	2071
Qy	6516	CCGTGCTGGAACCTTCCGCGGCCCAACCAAGTACCTCTACCTCCGCCCCCGAACATGTTG	6575
Dp	2070	CCGTGCTGGAACCTTCCGCGGCCCAACCAAGTACCTCTACCTCCGCCCCCGAACATGTTG	2011
Qy	6576	CCCCGGCCCGTGGCTCCCCCAAGCGAAGGGGCAAGAGGTCTCCAGAGCCAAACAGA	6635
Dp	2010	CCCCGGCCCGTGGCTCCCCCAAGCGAAGGGGCAAGAGGTCTCCAGAGCCAAACAGA	1951
Qy	6636	CGTGGGTCTTGGGTGGTGGTGAAGGACGGTATTGAACCTGTGTCCCCACCGAGGGGCTTGA	6695
Dp	1950	CGTGGGTCTTGGGTGGTGGTGAAGGACGGTATTGAACCTGTGTCCCCACCGAGGGGCTTGA	1891
Qy	6696	CGAGGCCAGAGGACATCCCGGAGTGTCTGTATACCCGCTGCTGTACCGGATATGGGAAACAGA	6755
Dp	1890	CGAGGCCAGAGGACATCCCGGAGTGTCTGTATACCCGCTGCTGTACCGGATATGGGAAACAGA	1831
Qy	6756	CGAGGCCAGAGGATGGGTCCAAAGTCTCCAGGCAACACAGCCAGCCGCGACGCTTCT	6815
Dp	1830	CGAGGCCAGAGGATGGGTCCAAAGTCTCCAGGCAACACAGCCAGCCGCGACGCTTCT	1771
Qy	6816	TCAGCAAGCTGACCGAGAGCACTCCGCGCATGTCTAAAGTCCAAAGAACAGAAATCAACA	6875
Dp	1770	TCAGCAAGCTGACCGAGAGCACTCCGCGCATGTCTAAAGTCCAAAGAACAGAAATCAACA	1711
Qy	6876	AGAAAGTGAACACCCCAACACCGGAATGAGGCTGAATCAATATATCAGCCAGCTTGGACGG	6935
Dp	1710	AGAAAGTGAACACCCCAACACCGGAATGAGGCTGAATCAATATATCAGCCAGCTTGGACGG	1651
Qy	6936	AGATCTTCAATATATGCTCCGCGCATCAACCGGAAACAGGCTTATATGACCTATATGAAGCCAGGCG	6995
Dp	1650	AGATCTTCAATATATGCTCCGCGCATCAACCGGAAACAGGCTTATATGACCTATATGAAGCCAGGCG	1591
Qy	6996	TGCAGGAACATGCGACGACCAACATGGGGGCTGAGAGGCGCATATATTAAGAAAGCACTATGG	7055
Dp	1590	TGCAGGAACATGCGACGACCAACATGGGGGCTGAGAGGCGCATATATTAAGAAAGCACTATGG	1531
Qy	7056	GTAAATATGACCAATGGGAAGAATCCCCCGCGCTCAGACGCGCAATGCTTTTAACCTCTGA	7115
Dp	1530	GTAAATATGACCAATGGGAAGAATCCCCCGCGCTCAGACGCGCAATGCTTTTAACCTCTGA	1471
Qy	7116	ATGCAATGACCAAGCTCCCGCTGCTATATGCGCATTAACCGCTGCTGACGGAAGTAGTACC	7175
Dp	1470	ATGCAATGACCAAGCTCCCGCTGCTATATGCGCATTAACCGCTGCTGACGGAAGTAGTACC	1411
Qy	7176	ACAACACTCACCTCGCCAGGTGGCGGCGGGAAGGCAAGGTCTCTGGAGAACCCAGCAGCC	7235
Dp	1410	ACAACACTCACCTCGCCAGGTGGCGGCGGGAAGGCAAGGTCTCTGGAGAACCCAGCAGCC	1351
Qy	7236	GAAGAGCAAGTCCCGCGGCCCTGGGCTATCTGGGAGACCGGCAACCTCTGTCTCTCT	7295
Dp	1350	GAAGAGCAAGTCCCGCGGCCCTGGGCTATCTGGGAGACCGGCAACCTCTGTCTCTCT	1291
Qy	7296	CAGTGCACCTCGAGAGGAGACTGCAACCGCCGGAACGCGCTCAACCAACCGCGTGTGGAGG	7355
Dp	1290	CAGTGCACCTCGAGAGGAGACTGCAACCGCCGGAACGCGCTCAACCAACCGCGTGTGGAGG	1231
Qy	7356	ACAAGCCCTGTCGCGAGTTTCAACGCAATTTCCCTTAACAACCCCTGATCATATCGGCTGC	7415





Db 541 CGAAGATCAACATGTGTAGAGCAGACATCTCTTAAGCTGAAAGAAAGCACAACAGCTG 600  
Qy 602 GAGGAGGAGGCTGCGCAAGCGCCGAGCCTGAGAGCCGGTGTACCGCGGCCATCGAG 661  
Db 601 GAGGAGGAGGCTGCGCAAGCGCCGAGCCTGAGAGCCGGTGTACCGCGGCCATCGAG 660  
Qy 662 TCGAAGCACCGGAGCCTGTGTGAGATCATCTACGAGAGAACCGGAAGAAAGGCTGAGGCT 721  
Db 661 TCGAAGCACCGGAGCCTGTGTGAGATCATCTACGAGAGAACCGGAAGAAAGGCTGAGGCT 720  
Qy 722 GCAATCGGATTTCTGGAAGGCTGTGGGCCCCAGGTGAGCTGCGCTGTACAAACGAGCC 781  
Db 721 GCAATCGGATTTCTGGAAGGCTGTGGGCCCCAGGTGAGCTGCGCTGTACAAACGAGCC 780  
Qy 782 TCCGACACCCGGAGATATATAGAAATCAAAATTAACGAGCGATGCGGAAGAGCTTA 841  
Db 781 TCCGACACCCGGAGATATATAGAAATCAAAATTAACGAGCGATGCGGAAGAGCTTA 840  
Qy 842 ATCTTGATCTTCAAGAGAGAAATCAGCTCGGAAAACAATGGAAGAGAAATTTCTGAG 901  
Db 841 ATCTTGATCTTCAAGAGAGAAATCAGCTCGGAAAACAATGGAAGAGAAATTTCTGAG 900  
Qy 902 CGCTATGACAGCTCATGGAAGGCTTTGAAAAAAAGGTGAGCGCATCGAAAAACAACCG 961  
Db 901 CGCTATGACAGCTCATGGAAGGCTTTGAAAAAAAGGTGAGCGCATCGAAAAACAACCG 960  
Qy 962 CGCCGCGCGGCGCAAGAGAGCAAGGTGCGCGAGTACTAGAAAAAGAGTTCCCTGAGATC 1021  
Db 961 CGCCGCGCGGCGCAAGAGAGCAAGGTGCGCGAGTACTAGAAAAAGAGTTCCCTGAGATC 1020  
Qy 1022 CGCAAGCAGCGGAGCTGAGAGACGATGACAGAGGAGGTGGCGCAGCGGGCAATGGG 1081  
Db 1021 CGCAAGCAGCGGAGCTGAGAGAGCGCATGCAAGAGGAGGTGGCGCAGCGGGCAATGGG 1080  
Qy 1082 CTGTCTATGTGCGCGCGCCGCAAGCAGACAAGAGTGTCAAGATCATGATGGCTCTCA 1141  
Db 1081 CTGTCTATGTGCGCGCGCCGCAAGCAGACAAGAGTGTCAAGATCATGATGGCTCTCA 1140  
Qy 1142 GAGCAGAGAACTTGGAGAAAGCAGATGCGCAGCTGGCCGTGATCCCCGCCATGCTGTAC 1201  
Db 1141 GAGCAGAGAACTTGGAGAAAGCAGATGCGCAGCTGGCCGTGATCCCCGCCATGCTGTAC 1200  
Qy 1202 GACGCTGACAGAGGCGCATCAAGTTCATCAACATGAAGGCGCTTATGGCCGACCCCATG 1261  
Db 1201 GACGCTGACAGAGGCGCATCAAGTTCATCAACATGAAGGCGCTTATGGCCGACCCCATG 1260  
Qy 1262 AAGGTGTACAAAGACCGCCAGGTCAATGAACATGTGAGTGAAGCAGAGAAAGACCTTC 1321  
Db 1261 AAGGTGTACAAAGACCGCCAGGTCAATGAACATGTGAGTGAAGCAGAGAAAGACCTTC 1320  
Qy 1322 CGGAGAAATTCAATGACATATCCCAAGAACTTTGGCTGTATGCAATTTCTGAGAGAG 1381  
Db 1321 CGGAGAAATTCAATGACATATCCCAAGAACTTTGGCTGTATGCAATTTCTGAGAGAG 1380  
Qy 1382 AAGACAGTGGCTGAGTGGCGCTCTATTCTACTACCTTAAGAAAGTGAAGAACTTAAG 1441  
Db 1381 AAGACAGTGGCTGAGTGGCGCTCTATTCTACTACTTAAGAAAGTGAAGAACTTAAG 1440  
Qy 1442 AAGCTTGTGAGACGGAAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACGACGAG 1501  
Db 1441 AAGCTTGTGAGACGGAAGCTATCGGCGCCGCGGCAAGAGCCAGCAGCAACAACGACGAG 1500  
Qy 1502 CAG 1561  
Db 1501 CAG 1560  
Qy 1562 GATGAGAAAG 1621  
Db 1561 GATGAGAAAG 1620  
Qy 1622 GACAAAGAAAG 1681  
Db 1621 GACAAAGAAAG 1680

Qy 1682 AAGAGGCTGTGGCTTCCAAAGGCGCAAAACTGCCAAAGAGCCAGGGAAGACGCAAAAGC 1741  
Db 1681 AAGAGGCTGTGGCTTCCAAAGGCGCAAAACTGCCAAAGAGCCAGGGAAGAGCAAAAGC 1740  
Qy 1742 CGCATCACCGGCTCAATGGCTTAATGAGCCCAACGAGAGAGGCTCATCCCCGACGAG 1801  
Db 1741 CGCATCACCGGCTCAATGGCTTAATGAGCCCAACGAGAGAGGCTCATCCCCGACGAG 1800  
Qy 1802 AGCGCCGAGCTGGCTCCATGAGAGCTGAATGAGAGTCTCGCTGACAGAAAGAAATG 1861  
Db 1801 AGCGCCGAGCTGGCTCCATGAGAGCTGAATGAGAGTCTCGCTGACAGAAAGAAATG 1860  
Qy 1862 GAAACAGCAAGAAAGAGTCTCTGGAACAGCGCCGCAAGCTGTGCGCATTCGCCGAGATG 1921  
Db 1861 GAAACAGCAAGAAAGAGTCTCTGGAACAGCGCCGCAAGCTGTGCGCATTCGCCGAGATG 1920  
Qy 1922 GTGGGCTCCAGACTGTGTGCAGTGTAAAGAACTTCTACTTCACTTAAGAAAGGACGAG 1981  
Db 1921 GTGGGCTCCAGACTGTGTGCAGTGTAAAGAACTTCTACTTCACTTAAGAAAGGACGAG 1980  
Qy 1982 AACCTGATGAGATCTTGAAGAGCACAAGCTGAAGATGGAAGAGAGAGAGAGAGAGAG 2041  
Db 1981 AACCTGATGAGATCTTGAAGAGCACAAGCTGAAGATGGAAGAGAGAGAGAGAGAGAG 2040  
Qy 2042 AGGAAGAAAGAAAGCGCGCGCGCGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2101  
Db 2041 AGGAAGAAAGAAAGCGCGCGCGCGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100  
Qy 2102 GAGGATGAGAGATGAGAGGCGGTGCGGCGGTGAGCGGAATGAGAGAGAGATGTGTAGAGAG 2161  
Db 2101 GAGGATGAGAGATGAGAGGCGGTGCGGCGGTGAGCGGAATGAGAGAGAGATGTGTAGAGAG 2160  
Qy 2162 GCTGAAGCTTTCATATCCCTCTGTGGAATGAGGTGCCCAAGAGGGAATGCAATGTGCCCAAGC 2221  
Db 2161 GCTGAAGCTTTCATATCCCTCTGTGGAATGAGGTGCCCAAGAGGGAATGCAATGTGCCCAAGC 2220  
Qy 2222 ACTGTCAACAAAGCTCAGACACCGAGACATCCCTCTCTTCAACTGAGGCGGCGCAAG 2281  
Db 2221 ACTGTCAACAAAGCTCAGACACCGAGACATCCCTCTCTTCAACTGAGGCGGCGCAAG 2280  
Qy 2282 GACACAGGCAAGATGAGGCGCCAGAGCCCGCAGCACTTGGGCGCGAGCGGCCACCCCA 2341  
Db 2281 GACACAGGCAAGATGAGGCGCCAGAGCCCGCAGCACTTGGGCGCGAGCGGCCACCCCA 2340  
Qy 2342 GGGCCACCCACCCCAACAGAGAGACATCCCGGAGCCCAATGAGGCCACCCGGGCTCT 2401  
Db 2341 GGGCCACCCACCCCAACAGAGAGACATCCCGGAGCCCAATGAGGCCACCCGGGCTCT 2400  
Qy 2402 GAAAGCCAGGAGCCCTTACGCGCCCGCAGACAGACCCCGCATGCGCTCTGCACTCTCTCT 2461  
Db 2401 GAAAGCCAGGAGCCCTTACGCGCCCGCAGACAGACCCCGCATGCGCTCTGCACTCTCTCTCT 2460  
Qy 2462 GTGTGTCCCAAG 2521  
Db 2461 GTGTGTCCCAAG 2520  
Qy 2522 GAGAGACAGAGAGCCCGCGCGCTGAGAGCTGTGACAGTGTGACACAGAGAGAGAGAGAGAG 2581  
Db 2521 GAGAGACAGAGAGCCCGCGCGCTGAGAGCTGTGACAGTGTGACACAGAGAGAGAGAGAGAG 2580  
Qy 2582 CCGGTCAAGAGAGAGTGCAGCGAGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2641  
Db 2581 CCGGTCAAGAGAGAGTGCAGCGAGAGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
Qy 2642 GAGGCGCTGAGAGCCAGCGCCAGAGGAGGCGCTCAAGCAGAGAGAGAGAGAGAGAGAGAG 2701  
Db 2641 GAGGCGCTGAGAGCCAGCGCCAGAGGAGGCGCTCAAGCAGAGAGAGAGAGAGAGAGAGAG 2700  
Qy 2702 GGCAGGAGCCACATGTGCAAGAGCTTGGGCGCCCGCAGAGACAGCGCATCTCAATGTCTAAC 2761  
Db 2701 GGCAGGAGCCACATGTGCAAGAGCTTGGGCGCCCGCAGAGACAGCGCATCTCAATGTCTAAC 2760

QY 2762 TCAGTCGACGAGGTGATGAGGCCGAGCGGCGGACAAACCGGGCTGTGTCCCA 2821  
DB 2761 TGCAGTGCAGAGAGGTGATGAGGCCGAGCGGCGGACAAAGACCGGCTGTGTCCCA 2820  
QY 2822 AGGCCAGGCTCTCTACCCCGA CTGGCGACCCCGGGGCAATGGCTCACTCCGAGAAGCCA 2881  
DB 2821 AGGCCAGGCTCTCTACCCCGA CTGGCGACCCCGGGGCAATGGCTCTCACTCCGAGAAGCCA 2880  
QY 2882 CTGACCTTGAAAGCACTGAAAGCAGCGAGCGGTGCATCTCCCGCACTCCAGGTCACTAA 2941  
DB 2881 CTGACCTTGAAAGCACTGAAAGCAGCGAGCGGTGCATCTCCCGCACTCCAGGTCACTAA 2940  
QY 2942 GTTCATGAGCCCCCGGGAGAGCGAGTCCGACCAAGCAGCTCTCCAGCCCCCAGC 3001  
DB 2941 GTTCATGAGCCCCCGGGAGAGCGAGTCCGACCAAGCAGCTCTCCAGCCCCCAGC 3000  
QY 3002 CCACGCGCAAAACCTGACAGCGGAGCGAGCGGCCCTCAGCAGCTGGAGCAGACCCCGG 3061  
DB 3001 CCACGCGCAAAACCTGACAGCGGAGCGAGCGGCCCTCAGCAGCTGGAGCAGACCCCGG 3060  
QY 3062 GGCAGAGAGAGGCGCGGACCCCGCGCGACAAAGAGGCTTGCAGCGCGAGGCCAG 3121  
DB 3061 GGCAGAGAGAGGCGCGGACCCCGCGCGACAAAGAGGCTTGCAGCGCGAGGCCAG 3120  
QY 3122 AAGCTGCTGGGGAACCCCTTGCTGGAATTCCGGCTGTGCCCTTCCCGGTGCCCGGT 3181  
DB 3121 AAGCTGCTGGGGAACCCCTTGCTGGAATTCCGGCTGTGCCCTTCCCGGTGCCCGGT 3180  
QY 3182 GAGGTGATCAAGGCTCCCCCGATGCGCGGACCCCTCAGCCTTCCCTCAGGCTCACT 3241  
DB 3181 GAGGTGATCAAGGCTCCCCCGATGCGCGGACCCCTCAGCCTTCCCTCAGGCTCACT 3240  
QY 3242 GGTCAACCACTGCGCCCTGGGCTCTCATGACACTGCGCGGCGGCTCTGCGCGGCCACCC 3301  
DB 3241 GGTCAACCACTGCGCCCTGGGCTCTCATGACACTGCGCGGCGGCTCTGCGCGGCCACCC 3300  
QY 3302 ACCATCTCGAACCCGCTCTCTCTCTCTCTGCGCAAGACCCCAAGCTCTCGAAGG 3361  
DB 3301 ACCATCTCGAACCCGCTCTCTCTCTCTCTGCGCAAGACCCCAAGCTCTCGAAGG 3360  
QY 3362 CAATATAGGTGCATCTCCCAAGAAATGTGATCAGCTCAGCTCCGCTCAGAGCAT 3421  
DB 3361 CAATATAGGTGCATCTCCCAAGAAATGTGATCAGCTCAGCTCCGCTCAGAGCAT 3420  
QY 3422 GCCAAGGCGCGGCTGGCTGTGACCATGAGGCTGCGGCTGCCCATGAGACCCCAAAAG 3481  
DB 3421 GCCAAGGCGCGGCTGGCTGTGACCATGAGGCTGCGGCTGCCCATGAGACCCCAAAAG 3480  
QY 3482 CTGGCACCCTTTCAGCGAGTGAAGCAGAGCAGCTGTCCCAAGGCTGCTGAGGCTGAGGCTG 3541  
DB 3481 CTGGCACCCTTTCAGCGAGTGAAGCAGAGCAGCTGTCCCAAGGCTGCTGAGGCTGAGGCTG 3540  
QY 3542 CCGAGAGCGTGGGGGTGCCACAGCCCAAGAGGGTCCGTGCTGAGAGGAGCAGCTCG 3601  
DB 3541 CCGAGAGCGTGGGGGTGCCACAGCCCAAGAGGGTCCGTGCTGAGAGGAGCAGCTCG 3600  
QY 3602 GGCCTAGTTCCGGGCGGAAGCATCACTAAAGGCAATCCAGACACAGGCTGCTCGAC 3661  
DB 3601 GGCCTAGTTCCGGGCGGAAGCATCACTAAAGGCAATCCAGACACAGGCTGCTCGAC 3660  
QY 3662 AGCGCATCATATACCGGGGCTTCATCACTCCAGGCAAGCAGCTGACCTCTGTAAAG 3720  
DB 3661 AGCGCATCATATACCGGGGCTTCATCACTCCAGGCAAGCAGCTGACCTCTGTAAAG 3720  
QY 3722 GGCACATCATCAGAGATATCGGCGAGGACAGCCGAGTGTGAGACCGCGCGGAG 3781  
DB 3721 GGCACATCATCAGAGATATCGGCGAGGACAGCCGAGTGTGAGACCGCGCGGAG 3780  
QY 3782 GACAGCTTGCCCAAGGCGCAGCTCATCTTCAAGAGCAAGAGGCGCAGCTTGTCTAT 3841  
DB 3781 GACAGCTTGCCCAAGGCGCAGCTCATCTTCAAGAGCAAGAGGCGCAGCTTGTCTAT 3840  
QY 3842 GAGGTGCGATGTCTGTGACCAAGTGTCCAAAGAGAGCGGCAAGACAGCTCAGAGACC 3901

DB 3841 GAGGTGCGATGTCTGTGACCAAGTGTCCAAAGAGAGCGGCAAGCAGCTCAGAGACC 3900  
QY 3902 CCCCATGAGACGGCGGCGCCCAAGGCACTTATGACATGATGAGGGCGCGCTGGGAG 3961  
DB 3901 CCCCATGAGACGGCGGCGCCCAAGGCACTTATGACATGATGAGGGCGCGCTGGGAG 3960  
QY 3962 GCCATCTCTCAGCAGCATTCGAGGTCTCATGGGCGGTGCATCTCCGCGAGCGACAC 4021  
DB 3961 GCCATCTCTCAGCAGCATTCGAGGTCTCATGGGCGGTGCATCTCCGCGAGCGACAC 4020  
QY 4022 AGCCCCACACCTCAAGAGAGAGACCAATCCGCGGTTCATACCAAGGATTCCT 4081  
DB 4021 AGCCCCACACCTCAAGAGAGAGACCAATCCGCGGTTCATACCAAGGATTCCT 4080  
QY 4082 CGGTCTTACGTGAGGAGCAGAGAGACTTACCTGCTGGGAGGCAAGCTCTTAAAGCG 4141  
DB 4081 CGGTCTTACGTGAGGAGCAGAGAGACTTACCTGCTGGGAGGCAAGCTCTTAAAGCG 4140  
QY 4142 GAGGCAAGGCTCGGCGGCGGCTCAAGGAGCTTCAAGGAGCTTCAAGAGCGAG 4201  
DB 4141 GAGGCAAGGCTCGGCGGCGGCTCAAGGAGCTTCAAGGAGCTTCAAGAGCGAG 4200  
QY 4202 GCGCTGGGCGGCTTGAAGTGAAGCGGCGGCTTGAAGGCGGCTTGAAGGAG 4261  
DB 4201 GCGCTGGGCGGCTTGAAGTGAAGCGGCGGCTTGAAGGCGGCTTGAAGGAG 4260  
QY 4262 GCGGCGGCTCTCATGAGATCCGCGGAGAGAGCTGCGGCAACGCGCGAGCTGCC 4321  
DB 4261 GCGGCGGCTCTCATGAGATCCGCGGAGAGAGCTGCGGCAACGCGCGAGCTGCC 4320  
QY 4322 CTGGCGCGGCGGCTCAAGAGAGGCTCATCAGCAGAGGACCCGCTCAAGTACAC 4381  
DB 4321 CTGGCGCGGCGGCTCAAGAGAGGCTCATCAGCAGAGGACCCGCTCAAGTACAC 4380  
QY 4382 ACCGGGCGTCCACACCTGCTCCAAAGACGAGCTTCCCTCATGCGGCGGCC 4441  
DB 4381 ACCGGGCGTCCACACCTGCTCCAAAGACGAGCTTCCCTCATGCGGCGGCC 4440  
QY 4442 GGCAGGAGTTCACACCGGTCACCGGCTGATGATGAGCGGCGGCGCATGAG 4501  
DB 4441 GGCAGGAGTTCACACCGGTCACCGGCTGATGATGAGCGGCGGCGCATGAG 4500  
QY 4502 CTGTCTCTCTACGAGAGAGCTTGAAGCGGCGGCGGAGCGGCGGAGCTCGGGGCG 4561  
DB 4501 CTGTCTCTCTACGAGAGAGCTTGAAGCGGCGGCGGAGCGGCGGAGCTCGGGGCG 4560  
QY 4562 TCCATTTGGCGGCGGCGGCTGTGATTTGTGCTGAGCTGGGTAACCGCGGAGAGCCCC 4621  
DB 4561 TCCATTTGGCGGCGGCGGCTGTGATTTGTGCTGAGCTGGGTAACCGCGGAGAGCCCC 4620  
QY 4622 CTGACCTATGAGAGCAGCGGAGCCTTTTGGCGGCGACCTTCCAGAGAGTTGCGCGGT 4681  
DB 4621 CTGACCTATGAGAGCAGCGGAGCCTTTTGGCGGCGACCTTCCAGAGAGTTGCGCGGT 4680  
QY 4682 ACCATGCGGAGCGGCGGCTGTGAGAGAGGAGCCTTTCTGTCAAGCAGGATCC 4741  
DB 4681 ACCATGCGGAGCGGCGGCTGTGAGAGAGGAGCCTTTCTGTCAAGCAGGATCC 4740  
QY 4742 CAGGACCGAAAGCTGACGTGACGCTGTGTGAGATTCGCAAGTCCCGGCAAGCAGCTG 4801  
DB 4741 CAGGACCGAAAGCTGACGTGACGCTGTGTGAGATTCGCAAGTCCCGGCAAGCAGCTG 4800  
QY 4802 CCGGAGCAACCCCAACCCCATCTGCGCTATGAGCAGCTGCTTGGGGGCGTGAAGTGC 4861  
DB 4801 CCGGAGCAACCCCAACCCCATCTGCGCTATGAGCAGCTGCTTGGGGGCGTGAAGTGC 4860  
QY 4862 GTGACCTGTATTCGAGGCAATCCCGCTTGCATTCAGACCCCACTTCATACCCCGGCG 4921  
DB 4861 GTGACCTGTATTCGAGGCAATCCCGCTTGCATTCAGACCCCACTTCATACCCCGGCG 4920  
QY 4922 ATCCCTCTGAGAGCGGCGGCTGTACTGCTGCCCGACACTGGCGGCCCAACCCACCC 4981

Db 4921 ATCCCTGTGAGCAGCCGCTGCTACTACCTGCCCCGACACTGCGCCCCAAACCCACC 4980  
Qy 4982 TACCGCACCCTGTACCCACCTACTCATCCGCGGCTAACCCGACAGCGCGGCTGAG 5041  
Db 4981 TACCGCACCCTGTACCCACCTACTCATCCGCGGCTAACCCGACAGCGCGGCTGAG 5040  
Qy 5042 AACCGGACAGCATCATATGATCATCACTCGCAGAGATGACCAACACGCGC 5101  
Db 5041 AACCGGACAGCATCATATGATCATCACTCGCAGAGATGACCAACACGCGC 5100  
Qy 5102 ACCGCAATGAGCCAGAGCTGATATGCTGAGGGGCTCTCGCCCCGAGTCTTGCTG 5161  
Db 5101 ACCGCAATGAGCCAGAGCTGATATGCTGAGGGGCTCTCGCCCCGAGTCTTGCTG 5160  
Qy 5162 GCACCTCACTACGCTGCGGGTCCCGAGGACATGACCTGTGCCAAGTGCACACTG 5221  
Db 5161 GCACCTCACTACGCTGCGGGTCCCGAGGACATGACCTGTGCCAAGTGCACACTG 5220  
Qy 5222 CCTGTGCTGTGCCCCCGACACCAAGGACCCGACGACCGGCATGACCGGCTTGCTAC 5281  
Db 5221 CCTGTGCTGTGCCCCCGACACCAAGGACCCGACGACCGGCATGACCGGCTTGCTAC 5280  
Qy 5282 CTGCCACCGCGCCCCAGCCCTTCAAGAGCCGCGCACAGAGCTCCCACTTCTCCAGAG 5341  
Db 5281 CTGCCACCGCGCCCCAGCCCTTCAAGAGCCGCGCACAGAGCTCCCACTTCTCCAGAG 5340  
Qy 5342 GGTCCACACACTTGAACAAACCAACACACAGTCTCTGTCGAGGGGAGCGAGACCG 5401  
Db 5341 GGTCCACACACTTGAACAAACCAACACACAGTCTCTGTCGAGGGGAGCGAGACCG 5400  
Qy 5402 GATCGAGACGGGACCGGGATCGGGAGCGGGAAAGTCCATCTCTCACTCAACAGAG 5461  
Db 5401 GATCGAGACGGGACCGGGATCGGGAGCGGGAAAGTCCATCTCTCACTCAACAGAG 5460  
Qy 5462 GTGAGACACGACCCATCTGAGACCTGTGTACAGACCAAGAGCGGACAGCGGACG 5521  
Db 5461 GTGAGACACGACCCATCTGAGACCTGTGTACAGACCAAGAGCGGACAGCGGACG 5520  
Qy 5522 AACGGGGGGGTGGGGGACAGACCGCCCCCGCTCCCATCTCCCATGCCACAGAC 5581  
Db 5521 AACGGGGGGGTGGGGGACAGACCGCCCCCGCTCCCATCTCCCATGCCACAGAC 5580  
Qy 5582 TCGCCCATCTCCCTGAGACCCAGATGACCTTCAGACAGAGACCCAGTGTCTTCAAC 5641  
Db 5581 TCGCCCATCTCCCTGAGACCCAGATGACCTTCAGACAGAGACCCAGTGTCTTCAAC 5640  
Qy 5642 ACAGGATGAAGGATATCATCAACGCTGTGAGACCCAGACCGACCGGCTTGAAGTCC 5701  
Db 5641 ACAGGATGAAGGATATCATCAACGCTGTGAGACCCAGACCGGCTTGAAGTCC 5700  
Qy 5702 ACTTCACCTCTCTCAACCCGTTGGCCAGAGTGCACATTCACACTGACCACTGCCCCA 5761  
Db 5701 ACTTCACCTCTCTCAACCCGTTGGCCAGAGTGCACATTCACACTGACCACTGCCCCA 5760  
Qy 5762 CTGGGAGGACCTTCGATGGGGTCTACCCCTAACCTCATGAGGCGGCTTGTGAGCCAA 5821  
Db 5761 CTGGGAGGACCTTCGATGGGGTCTACCCCTAACCTCATGAGGCGGCTTGTGAGCCAA 5820  
Qy 5822 GAGGCCCCCGGGTGCACCGGCGAGAGCGGCCCGGACAGACACACCGGCAATGCTTCTC 5881  
Db 5821 GAGGCCCCCGGGTGCACCGGCGAGAGCGGCCCGGACAGACACACCGGCAATGCTTCTC 5880  
Qy 5882 GCCAAGCCCCCAGCCGCTTCGGGGCTGAGAGCCGCTCTCTCCCGCAGCAAGGGCTGAG 5941  
Db 5881 GCCAAGCCCCCAGCCGCTTCGGGGCTGAGAGCCGCTCTCTCCCGCAGCAAGGGCTGAG 5940  
Qy 5942 CCGCGGCCCCCTATGTGCTCTGTCTCTGAGCAGGACCATGCGCGCACCCCTGCGAAG 6001  
Db 5941 CCGCGGCCCCCTATGTGCTCTGTCTCTGAGCAGGACCATGCGCGCACCCCTGCGAAG 6000  
Qy 6002 AACCTGACCTTCAACAGCGACCGGAGCCCGCGGCGCACTTGTGAGCTTGTGAG 6061  
Db 6001 AACCTGACCTTCAACAGCGACCGGAGCCCGCGGCGCACTTGTGAGCTTGTGAG 6060

Qy 6062 CCGCACCGGGAAAAGACTCAAAAGTAAACCTTTTCATCCAGAACTGAACTCCGTTCT 6121  
Db 6061 CCGCACCGGGAAAAGACTCAAAAGTAAACCTTTTCATCCAGAACTGAACTCCGTTCT 6120  
Qy 6122 CTGGGTTACAGGACGACGCTACAGCCCCGAAAGGGGTGAGGCCCTGACGCTGTGAGC 6181  
Db 6121 CTGGGTTACAGGACGACGCTACAGCCCCGAAAGGGGTGAGGCCCTGACGCTGTGAGC 6180  
Qy 6182 TCACCCAGCTGACCCACGACAAAGGGGCTCCCAAGCCTGTGAGAGGCTCGACAGAGC 6241  
Db 6181 TCACCCAGCTGACCCACGACAAAGGGGCTCCCAAGCCTGTGAGAGGCTCGACAGAGC 6240  
Qy 6242 CACTTGAAGGGGAGGTGCGGCGCAAGACGAGGCCCGGTGAAGCTTGTGCGGGAGGCGC 6301  
Db 6241 CACTTGAAGGGGAGGTGCGGCGCAAGACGAGGCCCGGTGAAGCTTGTGCGGGAGGCGC 6300  
Qy 6302 GCCCACTTCCACACTTGTGCGCGCTGTCTGAGAGGCTGACGCTCTGTCAGCCGCTGCTC 6361  
Db 6301 GCCCACTTCCACACTTGTGCGCGCTGTCTGAGAGGCTGACGCTCTGTCAGCCGCTGCTC 6360  
Qy 6362 CAGACCGCCCCAGGGGTCAAGGTCAACAGCGGGGTGACCTGTGCGCAGACATCAAT 6421  
Db 6361 CAGACCGCCCCAGGGGTCAAGGTCAACAGCGGGGTGACCTGTGCGCAGACATCAAT 6420  
Qy 6422 GAGGTATCAACAAGGACTACACCGGACCAACCCACAGAGCTCAGCGCACCCCTGCGC 6481  
Db 6421 GAGGTATCAACAAGGACTACACCGGACCAACCCACAGAGCTCAGCGCACCCCTGCGC 6480  
Qy 6482 GCCCCCTCTACTCTTCTCTGCGGGCAGCTGCCCCGTCTGAGACTTCCGCGGCCACCC 6541  
Db 6481 GCCCCCTCTACTCTTCTCTGCGGGCAGCTGCCCCGTCTGAGACTTCCGCGGCCACCC 6540  
Qy 6542 AGTGAACCTTACCTCTCCGCCCCCGACATGTGTCCCCCGGCTGTGAGCTTCCCCACAGC 6601  
Db 6541 AGTGAACCTTACCTCTCCGCCCCCGACATGTGTCCCCCGGCTGTGAGCTTCCCCACAGC 6600  
Qy 6602 GAAAGGGGAGAGGCTCTCAGAGCAACAAGCTGAGTCTTGTGGTGTGAGTGAAGAC 6661  
Db 6601 GAAAGGGGAGAGGCTCTCAGAGCAACAAGCTGAGTCTTGTGGTGTGAGTGAAGAC 6660  
Qy 6662 GGTATTTGAACCTGTGTCCCAACCGAGGGCATGACGAGCCAGAGGACTTCCGGAAGTCTC 6721  
Db 6661 GGTATTTGAACCTGTGTCCCAACCGAGGGCATGACGAGCCAGAGGACTTCCGGAAGTCTC 6720  
Qy 6722 GTGTACCCGCTGTGTACCGGGATGGGGAAACAGACGGAACCCAGAGGATGGGCTTCAAG 6781  
Db 6721 GTGTACCCGCTGTGTACCGGGATGGGGAAACAGACGGAACCCAGAGGATGGGCTTCAAG 6780  
Qy 6782 TCTCCAGGCAACACAGCGGACCGGAGCTTCTTCAAGCAAGCTGACGAGAGCACTCC 6841  
Db 6781 TCTCCAGGCAACACAGCGGACCGGAGCTTCTTCAAGCAAGCTGACGAGAGCACTCC 6840  
Qy 6842 GCCATGTCAATTCAGAGAGAGAGATCAACAGAACTGAACACCAACCGGAAT 6901  
Db 6841 GCCATGTCAATTCAGAGAGAGAGATCAACAGAACTGAACACCAACCGGAAT 6900  
Qy 6902 GAGCTGAATTAATATATCAGCGAGCTGGGAGCGAGATCTTCAATATGCCCCCATACCC 6961  
Db 6901 GAGCTGAATTAATATATCAGCGAGCTGGGAGCGAGATCTTCAATATGCCCCCATACCC 6960  
Qy 6962 GGAACAGGCTTATGACCTATGAAGGACGAGGGGTGAGAGAACTATGCCAGACCAACATG 7021  
Db 6961 GGAACAGGCTTATGACCTATGAAGGACGAGGGGTGAGAGAACTATGCCAGACCAACATG 7020  
Qy 7022 GGGCTGAGAGGCAATATTGAAGAGCACTCATGGGTAAATATGACCAATGTGGAAGAGTCC 7081  
Db 7021 GGGCTGAGAGGCAATATTGAAGAGCACTCATGGGTAAATATGACCAATGTGGAAGAGTCC 7080  
Qy 7082 CCGCGGCTCAGGCGCAATGCTTTTAAACCTTGTGAATGCAAGTGCAGGCTGCGCGTGCCT 7141  
Db 7081 CCGCGGCTCAGGCGCAATGCTTTTAAACCTTGTGAATGCAAGTGCAGGCTGCGCGTGCCT 7140

QY 7142 ATGCCATTAACCGCTGCTGACGAGACGAGTGAACAACACTACCTCCGACAGTGGCGGC 7201  
DB 7141 ATGCCATTAACCGCTGCTGACGAGACGAGTGAACAACACTACCTCCGACAGTGGCGGC 7200  
QY 7202 GGGAAAGGCGAAGGCTCTGCGGACGACCGACCGAAAAAGCCAAGTCCCGCGGCCCGGCGC 7261  
DB 7201 GGGAAAGGCGAAGGCTCTGCGGACGACCGACCGAAAAAGCCAAGTCCCGCGGCCCGGCGC 7260  
QY 7262 CTGGCATCTGGGGAACCGGCAACCTCTGTCTCTCTCACTGACGAGGAGGAGACTGCAAC 7321  
DB 7261 CTGGCATCTGGGGAACCGGCAACCTCTGTCTCTCTCACTGACGAGGAGGAGACTGCAAC 7320  
QY 7322 CGCCGAGCGCGCTCAACCAACCGCGTGTGGAGAGACAGGCGCTCGTCCGACAGTTCCACG 7381  
DB 7321 CGCCGAGCGCGCTCAACCAACCGCGTGTGGAGAGACAGGCGCTCGTCCGACAGTTCCACG 7380  
QY 7382 CCATTCCCTTCAACACCCCTGATCATGAGGAGGAGAGGAGGAGTGTAGTATGGCTTCCCAACC 7441  
DB 7381 CCATTCCCTTCAACACCCCTGATCATGAGGAGGAGAGGAGGAGTGTAGTATGGCTTCCCAACC 7440  
QY 7442 CCACCGGAGCTTCCCGCGGAGCGGCGCCCTCGCTGCGCCCAACACGCTGGAGACGAG 7501  
DB 7441 CCACCGGAGCTTCCCGCGGAGCGGCGCCCTCGCTGCGCCCAACACGCTGGAGACGAG 7500  
QY 7502 GAGCCCAAGCCACTGCTCTGCTGCAATGCAACAATCTTCCGACAGGAGTGA 7555  
DB 7501 GAGCCCAAGCCACTGCTCTGCTGCAATGCAACAATCTTCCGACAGGAGTGA 7554

## RESULT 9

US-09-819-104A-3

; Sequence 3, Application US/09819104A

ABN

; Publication No. US20030027137A1

; GENERAL INFORMATION:

; APPLICANT: Chem. J. Don

; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES

; TITLE OF INVENTION: AND USES THEREFOR

; FILE REFERENCE: UMG-030

; CURRENT APPLICATION NUMBER: US/09/819,104A

; CURRENT FILING DATE: 2001-03-27

; PRIOR APPLICATION NUMBER: 60/193,138

; PRIOR FILING DATE: 2000-03-29

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 7521

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1) .. (7521)

US-09-819-104A-3

Query Match 85.6%; Score 7334.6; DB 3; Length 7521;

Best Local Similarity 98.5%; Pred. No. 0;

Matches 7463; Conservative 0; Mismatches 34; Indels 78; Gaps 3;

QY 2 ATGTGGGCTCCACACAGCTTGTGCAACAGAGTGAAGGAGCCACTAGAGCCCGTACCGG 61  
DB 1 ATGTGGGCTCCACACAGCTTGTGCAACAGAGTGAAGGAGCCACTAGAGCCCGTACCGG 60  
QY 62 CCCCAAGCCTTCTCTAACCAGTGAAGATGCGCCGAGCGACACGAGCGTGGGCTCTG 121  
DB 61 CCCCAAGCCTTCTCTAACCAGTGAAGATGCGCCGAGCGACACGAGCGTGGGCTCTG 120  
QY 122 GAGTACGAGCACTCTCCGCGCATATGCTCTCCACCTGTGCGCGGCTCTCATATCCAG 181  
DB 121 GAGTACGAGCACTCTCCGCGCATATGCTCTCCACCTGTGCGCGGCTCTCATATCCAG 180  
QY 182 CCCGAGCGGCGAGGCGCTCCTGCTGTGATTCAGGCTCGGGGAATGAAGTCCGAG 241  
DB 181 CCCGAGCGGCGAGGCGCTCCTGCTGTGATTCAGGCTCGGGGAATGAAGTCCGAG 240

QY 242 GAGCTTCAACTGCGGCGCAGAGTCCCACTATACCTGCGCGAGCTGGGAGTCAAGATG 301  
DB 241 GAGCTTCAACTGCGGCGCAGAGTCCCACTATACCTGCGCGAGCTGGGAGTCAAGATG 300  
QY 302 GAGTTCAATTGAAGAGAGCGGCTTCGCTAGAGCTGCTCGTACCGCTGCGGACCG 361  
DB 301 GAGTTCAATTGAAGAGAGCGGCTTCGCTAGAGCTGCTCGTACCGCTGCGGACCG 360  
QY 362 TCACCCCTGCTGCGCAGCGGCGCAGGCTGCGGAGTCTGAAGACCTTCAACAGACCGTAC 421  
DB 361 TCACCCCTGCTGCGCAGCGGCGCAGGCTGCGGAGTCTGAAGACCTTCAACAGACCGTAC 420  
QY 422 CTGACGCGCAAGCTGGAACCGGCTGTCTTCCCCAGCGCCCCCGGCACTGACCTTGA 481  
DB 421 CTGACGCGCAAGCTGGAACCGGCTGTCTTCCCCAGCGCCCCCGGCACTGACCTTGA 480  
QY 482 GAGCTGTGCGGCGCAGGCTGTCCAAAGGAGAGCTGATCCAGAAATCAATGACCGCTGAC 541  
DB 481 GAGCTGTGCGGCGCAGGCTGTCCAAAGGAGAGCTGATCCAGAAATCAATGACCGCTGAC 540  
QY 542 CGAGGATCACCATGTGTAGACAGCAGATCTTAAAGCTGAAGAGAGCAGCAAGCTG 601  
DB 541 CGAGGATCACCATGTGTAGACAGCAGATCTTAAAGCTGAAGAGAGCAGCAAGCTG 600  
QY 602 GAGGAGAGGCTGCGCAACCGCGCCGAGGCTGGAAGAGCGCGTGTCAACGCGCCATGAG 661  
DB 601 GAGGAGAGGCTGCGCAACCGCGCCGAGGCTGGAAGAGCGCGTGTCAACGCGCCATGAG 660  
QY 662 TCGAAGCAGCGGAGCTGTGATGATCATCTTACGACGAGAACCGGAGAGAGCTGAC 721  
DB 661 TCGAAGCAGCGGAGCTGTGATGATCATCTTACGACGAGAACCGGAGAGAGCTGAC 720  
QY 722 GCACATCGAATTTTGAAGGCTGTGGGCGCCGAGGTGAGCTGCGCTGTACACGAGCCC 781  
DB 721 GCACATCGAATTTTGAAGGCTGTGGGCGCCGAGGTGAGCTGCGCTGTACACGAGCCC 780  
QY 782 TCCGACACCGGAGATATCATGAGAACATCAAAATTAACAGGCGATCGGAGAGAGCTA 841  
DB 781 TCCGACACCGGAGATATCATGAGAACATCAAAATTAACAGGCGATCGGAGAGAGCTA 840  
QY 842 ATCTTGTACTTCAAGAGAGAGATACGCTCGGAAACAATGGAAGCAGAGTTCTGCAG 901  
DB 841 ATCTTGTACTTCAAGAGAGAGATACGCTCGGAAACAATGGAAGCAGAGTTCTGCAG 900  
QY 902 CGCTATGACAGCTCATGAGAGCGCTTGGAAAAAAGGTGAGCGCATCGAAAAACAACCG 961  
DB 901 CGCTATGACAGCTCATGAGAGCGCTTGGAAAAAAGGTGAGCGCATCGAAAAACAACCG 960  
QY 962 CGCGCGCGGCGCAGGAGAGCAAGGTGCGCGAGTCTACGAAAGCAATGCTCTGAGATC 1021  
DB 961 CGCGCGCGGCGCAGGAGAGCAAGGTGCGCGAGTCTACGAAAGCAATGCTCTGAGATC 1020  
QY 1022 CGCAAGCAGCGGAGCTGACGAGAGCGCATGACAGAGCAGGAGTGGGCGACGCGGCGAGTGG 1081  
DB 1021 CGCAAGCAGCGGAGCTGACGAGAGCGCATGACAGAGCAGGAGTGGGCGACGCGGCGAGTGG 1077  
QY 1082 CTGTTCATGTGGCGCGCGCCGAGCGAGCAGAGGTTGTCAGAGATCATGATGCTCTCA 1141  
DB 1078 CTGTTCATGTGGCGCGCGCGCGAGCGAGCAGAGTGTCAAGATCATGATGCTCTCA 1137  
QY 1142 GAGCAGAGAGAACCTGGAAGAGATGCGCGACGCTGCGCGCTGATCCCGCCCATGCTGAC 1201  
DB 1138 GAGCAGAGAGAACCTGGAAGAGATGCGCGACGCTGCGCGCTGATCCCGCCCATGCTGAC 1197  
QY 1202 GACGCTGACAGCGGCGATCAAGTTTCAACATGAACGCGGCTTAAAGGCGAGCCCGCATG 1261  
DB 1198 GACGCTGACAGCGGCGATCAAGTTTCAACATGAACGCGGCTTAAAGGCGAGCCCGCATG 1257  
QY 1262 AAGGTGTACAAAGACCGCGCAGGTCAATGAACATGTGAGTGTGACAGAGAGAGACCTTC 1321  
DB 1258 AAGGTGTACAAAGACCGCGCAGGTCAATGAACATGTGAGTGTGAGAGAGAGAGACCTTC 1317  
QY 1322 CGGAGAAATTCATGACGATCCCAAGAACTTTGGCTGTATGCAATCTTCTGAGAGAG 1381

```
Db 1318 CGGAGAAAGTTTCATGACGATCCCAAGAACTTGGCCGTATCCGATCATCTTGAGAGG 1377
Qy 1382 AAGACAGTGCCTAAGTGCCTCTCTATTACTTACTGAAGTAAGTAAGCTTAAG 1441
Db 1378 AAGACAGTGCCTAAGTGCCTCTCTATTACTTACTGAAGTAAGTAAGCTTAAG 1437
Qy 1442 AGCTGTGTGAGACGAGACTATCGGCGCGCGGCAAGAGCAGACGAAACAACAGCAG 1501
Db 1438 AGCTGTGTGAGACGAGACTATCGGCGCGCGGCAAGAGCAGACGAAACAACAGCAG 1497
Qy 1502 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1561
Db 1498 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1557
Qy 1562 GATGAAAGAAAGAAAGAAAGAGCGGAGAAAGAAAGAAAGCGGAGGTGAGAAC 1621
Db 1558 GATGAAAGAAAGAAAGAAAGAGCGGAGAAAGAAAGCGGAGGTGAGAAC 1617
Qy 1622 GACAAAGAAAGAAAGCTCTCAAGAGAAAGACAGACGACTCAGGGAGAGCAACGAC 1681
Db 1618 GACAAAGAAAGAAAGCTCTCAAGAGAAAGACAGACGACTCAGGGAGAGCAACGAC 1677
Qy 1682 AAGAGAGCTGTGGCTCTCAAGAGCGCGCAAACTGCCAACGCGAGAGAGACGCAAG 1741
Db 1678 AAGAGAGCTGTGGCTCTCAAGAGCGCGCAAACTGCCAACGCGAGAGAGACGCAAG 1737
Qy 1742 CGGATCAACCGGCTCAATGGCTTAAGAGCCCAACAGCAGAGAGCGCATCACCCCG 1801
Db 1738 CGGATCAACCGGCTCAATGGCTTAAGAGCCCAACAGCAGAGAGCGCATCACCCCG 1797
Qy 1802 AGCGCGAGCTGGCTCAATGGAGCTGAATGAGATTTCTGCTGAGACAGAAAGAAAG 1861
Db 1798 AGCGCGAGCTGGCTCAATGGAGCTGAATGAGATTTCTGCTGAGACAGAAAGAAAG 1857
Qy 1862 GAAACAGCCAAAGAAAGTCTCTGAGAACACGCGCGCAACTGTGCGCATCGCCGAG 1921
Db 1858 GAAACAGCCAAAGAAAGTCTCTGAGAACACGCGCGCAACTGTGCGCATCGCCGAG 1917
Qy 1922 GTGGGCTCCAAAGCTGTGTGCGAGTGAAGACTTTCTTAACCTAACAAGAGAGCAG 1981
Db 1918 GTGGGCTCCAAAGCTGTGTGCGAGTGAAGACTTTCTTAACCTAACAAGAGAGCAG 1977
Qy 1982 AACCTGATGAGATCTTTGACGAGCACAAGCTGAAGTGAAGAGAGAGAGAGCGCG 2041
Db 1978 AACCTGATGAGATCTTTGACGAGCACAAGCTGAAGTGAAGAGAGAGAGAGCGCG 2037
Qy 2042 AGGAAAGAAAGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2101
Db 2038 AGGAAAGAAAGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2097
Qy 2102 GAGGATGAGAGATGAGAGCGGTGCGGCGTGAAGCGGAAATGAGAGAGATGTGAGAG 2161
Db 2098 GAGGATGAGAGATGAGAGCGGTGCGGCGTGAAGCGGAAATGAGAGAGATGTGAGAG 2157
Qy 2162 GCTGAAGCCTTACATGCTCTGCGGAATGAGGTGCCAGAGGGGAAATGCAGTGCCAG 2221
Db 2158 GCTGA-----AGCC 2166
Qy 2222 ACTGTCAACAACAGCTCAAGACACGAGAGATCCCTCTCTCACTGAGCGCGCAAG 2281
Db 2167 ACTGTCAACAACAGCTCAAGACACGAGAGATCCCTCTCTCACTGAGCGCGCAAG 2226
Qy 2282 GACACAGGCGAAGATGAGCGCCCAAGCGCCCAAGCGCTGCGCGCGCGCGCGCG 2341
Db 2227 GACACAGGCGAAGATGAGCGCCCAAGCGCCCAAGCGCTGCGCGCGCGCGCGCG 2286
Qy 2342 GCGCCACCAACCCCAACAGAGAGATCCCGGCGCGCGCGCGCGCGCGCGCGCGCT 2401
Db 2287 GCGCCACCAACCCCAACAGAGAGATCCCGGCGCGCGCGCGCGCGCGCGCGCGCT 2346
Qy 2402 GAAGCCACCGAGCGCTTACGCGCCCAACGACACCCCGCATCGCTCTGCACTCTCT 2461
Db 2347 GAAGCCACCTTACGCGCTTACGCGCCCAACGACACCCCGCATCTCTCTCTCTCT 2406
Qy 2462 GTGGTCCCCAAGAGAGAGAGAGAGAGACCGCAGCAGCGCCCCCAGTGAAGAGGG 2521
Db 2407 GTGGTCCCCAAGAGAGAGAGAGAGAGACCGCAGCAGCGCCCCCAGTGAAGAGGG 2466
Qy 2522 GAGGACAGAAAGCCCCCGGCTGAGAGCTGGCAATGAGCAACAGGAAAGCCGAGAG 2581
Db 2467 GAGGACAGAAAGCCCCCGGCTGAGAGCTGGCAATGAGCAACAGGAAAGCCGAGAG 2526
Qy 2582 CCCGTCAAGAGCGATGACAGAGAGAGCCGAGAGAGGGCGCGCCAAAGGACGCG 2641
Db 2527 CCCGTCAAGAGCGATGACAGAGAGAGCCGAGAGAGGGCGCGCCAAAGGACGCG 2586
Qy 2642 GAGGCGCTGAGGCGCAGCGCGAGGGGGGCTCAAGGCAAGAAAGAGAGGGCGGAG 2701
Db 2587 GAGGCGCTGAGGCGCAGCGCGAGGGGGGCTCAAGGCAAGAAAGAGGGCGGAG 2646
Qy 2702 GGCAGGGCCACCACTGCGCAAGAGCTCGGCGCGCCCCCAGAGACGCACTTCACTG 2761
Db 2647 GGCAGGGCCACCACTGCGCAAGAGCTCGGCGCGCCCCCAGAGACGCACTTCACTG 2706
Qy 2762 TGCAGTGCAGACGAGGTGATGAGGCTCGAGGCGCGCAAGAACCGGCTGTCTCC 2821
Db 2707 TGCAGTGCAGACGAGGTGATGAGGCTCGAGGCGCGCAAGAACCGGCTGTCTCC 2766
Qy 2822 AGGCGCAGGCTCTCAACCGCGACTGGCGACCCCGGGCCCAATGCTCACCCGAAAG 2881
Db 2767 AGGCGCAGGCTCTCAACCGCGACTGGCGACCCCGGGCCCAATGCTCACCCGAAAG 2826
Qy 2882 CTGGAAGCTGAAGAGCTGAAGACGAGCGGCTGACATCCCGCCATCCAGAGTCA 2941
Db 2827 CTGGAAGCTGAAGAGCTGAAGACGAGCGGCTGACATCCCGCCATCCAGAGTCA 2886
Qy 2942 GTTCATGAGCCCCCGGGAGAGACGAGCTCCCAACAGCAGTCTCCCGACCG 3001
Db 2887 GTTCATGAGCCCCCGGGAGAGACGAGCTCCCAACAGCAGTCTCCCGACCG 2946
Qy 3002 CCACGCGCAAACTGAGCGCGAGCGAGCGACGCCCTCAAGCTGAGCAGACGCCCG 3061
Db 2947 CCACGCGCAAACTGAGCGCGAGCGAGCGACGCCCTCAAGCTGAGCAGACGCCCG 3006
Qy 3062 GGCAGAGCAGAGCGCGGCAACCCCGCGCAAGAGAG----- 3100
Db 3007 GGCAGAGCAGAGCGCGGCAACCCCGCGCAAGAGAGCGCTGTCTTC 3066
Qy 3101 ---GCTTGCAGCGCGGCGCAAGAGCTGCTGGGAGCTTGTGAGACTTCGCG 3157
Db 3067 CCAGCTTGCAGCGCGGCGCAAGAGCTGCTGGGAGCTTGTGAGACTTCGCG 3126
Qy 3158 CTGCGCTTCCCGTGGCCCCCGGTGAGTGAATCAAGGCTTCCCGGATGCCCGGAC 3217
Db 3127 CTGCGCTTCCCGTGGCCCCCGGTGAGTGAATCAAGGCTTCCCGGATGCCCGGAC 3186
Qy 3218 TCAGCTTCTCTCAAGCTGCACTGAGTCAACCACTGCTGCGGAGCTCATGACCTG 3277
Db 3187 TCAGCTTCTCTCAAGCTGCACTGAGTCAACCACTGCTGCGGAGCTCATGACCTG 3246
Qy 3278 CGGCGCGTCTGCGCGGCGCCACCACTGTCMAACCGGCTTCCCTCATCTCTGCG 3337
Db 3247 CGGCGCGTCTGCGCGGCGCCACCACTGTCMAACCGGCTTCCCTCATCTCTGCG 3306
Qy 3338 AAGCACCCCAAGGTCTTCAAGAGGCAAAATGAGTCCATCTCCCAAGAAATGTCG 3397
Db 3307 AAGCACCCCAAGGTCTTCAAGAGGCAAAATGAGTCCATCTCCCAAGAAATGTCG 3366
Qy 3398 CTCACGTCCCGTACTCAGAGATGCAAGGCGCGGAGGCGCTGTCACTAGGAGCTG 3457
Db 3367 CTCACGTCCCGTACTCAGAGATGCAAGGCGCGGAGGCGCTGTCACTAGGAGCTG 3426
Qy 3458 CCCCTGCGCATGAGACCCCAAAAGCTGAGACCTTCAAGCGAGTGAAGAGAGAGCT 3517
Db 3427 CCCCTGCGCATGAGACCCCAAAAGCTGAGACCTTCAAGCGAGTGAAGAGAGAGCT 3486
```

QY 3518 TTCCCAACGGGGCCAGGCTGGGCCAATCGGAGAGCTGGGGGTGCCACAGCCCAAGAGGCG 3577  
DB 3487 TTCCCAACGGGGCCAGGCTGGGCCAATCGGAGAGCTGGGGGTGCCACAGCCCAAGAGGCG 3546  
QY 3578 TTCCGCTGAGAGGGGACAGCTCGGGCTGAGTTCCGGGGGGAAGATCAACCAAGGCATT 3637  
DB 3547 TTCCGCTGAGAGGGGACAGCTCTGGGCTGAGTTCCGGGGGGAAGATCAACCAAGGCATT 3606  
QY 3638 CCACAGCAACGGGGTGCCTTCGACAGCGCCATCAATACCGCGGCTCCATCAACCGACGCG 3697  
DB 3607 CCACAGCAACGGGGTGCCTTCGACAGCGCCATCAATACCGCGGCTCCATCAACCGACGCG 3666  
QY 3698 ACCGCACTGACGTCTCTGTACAAAGGACCAATCAACAGGATCATGCGGAGAGACAGCCCG 3757  
DB 3667 ACCGCACTGACGTCTCTGTACAAAGGACCAATCAACAGGATCATGCGGAGAGACAGCCCG 3726  
QY 3758 AGTCGCTTGGACCGGGGCGGGGAGAGACAGCTGGCCCAAGGGCAAGTCACTTCAAGAGCG 3817  
DB 3727 AGTCGCTTGGACCGGGGCGGGGAGAGACAGCTGGCCCAAGGGCAAGTCACTTCAAGAGCG 3786  
QY 3818 AAGAAAGGACAGTCTTGTCTATGAGGGTGGCATGTCTGTGACCAAGTCTCAAGAGAG 3877  
DB 3787 AAGAAAGGACAGTCTTGTCTATGAGGGTGGCATGTCTGTGACCAAGTCTCAAGAGAG 3846  
QY 3878 GACGGCAAGAGAGCTCAAGAACCCCCCATGAGACGGCGGCCCCCAAGCGCACTATGAC 3937  
DB 3847 GACGGCAAGAGAGCTCAAGAACCCCCCATGAGACGGCGGCCCCCAAGCGCACTATGAC 3906  
QY 3938 ATGATGAGAGGGCGGGTGGGAGAGCAATCTCTCAAGCCAGGATGGAAGTCTCAATGGGC 3997  
DB 3907 ATGATGAGAGGGCGGGTGGGAGAGCAATCTCTCAAGCCAGGATGGAAGTCTCAATGGGC 3966  
QY 3998 CGTGCATCTCCGCGGAGGAGACAGACCCCACTCAAAAGAGAGACCAATCCGC 4057  
DB 3967 CGTGCATCTCCGCGGAGGAGACAGACCCCACTCAAAAGAGAGACCAATCCGC 4026  
QY 4058 GGGTCATCAACAAGGAGTCCCTGGGTCTTACGTGAGAGCAACAGAGAGCTAACCTGCGT 4117  
DB 4027 GGGTCATCAACAAGGAGTCCCTGGGTCTTACGTGAGAGCAACAGAGAGCTAACCTGCGT 4086  
QY 4118 CGGGAGGCAAGGCTCTAAAGGGGAGGAGGAGCCCTCGGCCCAACCGGCTCAAGGAG 4177  
DB 4087 CGGGAGGCAAGGCTCTAAAGGGGAGGAGGAGCCCTCGGCCCAACCGGCTCAAGGAG 4146  
QY 4178 CTGACCGAGGCTTACAGAGCGAGGCCCTGGGCCCTCTGAAGCTGAAGCGGCCCATGAG 4237  
DB 4147 CTGACCGAGGCTTACAGAGCGAGGCCCTGGGCCCTCTGAAGCTGAAGCGGCCCATGAG 4206  
QY 4238 GGCCTGTGGCCACGATGAAGAGGGCGGGCGCTCCATCATGAGATCCCGCGAGAGAG 4297  
DB 4207 GGCCTGTGGCCACGATGAAGAGGGCGGGCGCTCCATCATGAGATCCCGCGAGAGAG 4266  
QY 4298 CTGGGGCAACGCGCCGAGCTGCGCTCTGGGCCCGGCGGCTCAAGAGGGCTCCATCAG 4357  
DB 4267 CTGGGGCAACGCGCCGAGCTGCGCTCTGGGCCCGGCGGCTCAAGAGGGCTCCATCAG 4326  
QY 4358 CAGGCAACCCCGCTCAAGTACAGACACCGGCGGTCAACCATGTGCTCCAAAAGACAGAC 4417  
DB 4327 CAGGCAACCCCGCTCAAGTACAGACACCGGCGGTCAACCATGTGCTCCAAAAGACAGAC 4386  
QY 4418 GTACGCTCTCTCATGTGGAGCGCCGAGCGGAGGTTCCCAACCGGTGATGATG 4477  
DB 4387 GTACGCTCTCTCATGTGGAGCGCCGAGCGGAGGTTCCCAACCGGTGATGATG 4446  
QY 4478 ATGGCGAGACGCGCGGAGCACTGGAATGTGCTGTACAGAGAGAGCTTGAAGCGCGGCA 4537  
DB 4447 ATGGCGAGACGCGCGGAGCACTGGAATGTGCTGTACAGAGAGAGCTTGAAGCGCGGCA 4506  
QY 4538 GGGACCGACAGAGCTCGGGGGGCTCAATTGCGCGGGCGCCCGGTCAATTGTGCTGAG 4597  
DB 4507 GGGACCGACAGAGCTCGGGGGGCTCAATTGCGCGGGCGCCCGGTCAATTGTGCTGAG 4566

QY 4558 CTGGGTAAAGCGCGGCAAGACCCCTTGACCTATGAGACACAGGGGCAACCTTTGCGGCG 4657  
DB 4567 CTGGGTAAAGCGCGGCAAGACCCCTTGACCTATGAGAGACACAGGGGCAACCTTTGCGGCG 4626  
QY 4658 CACCTCCCAAGAGTTGCGCGGTGACCAATGGGGAGGCCACGCGGCGCTGAGAGAGGCG 4717  
DB 4627 CACCTCCCAAGAGTTGCGCGGTGACCAATGGGGAGGCCACGCGGCGCTGAGAGAGGCG 4686  
QY 4718 AGCTTTGTTCAGAGAGGATCCAGAGACGAAAAGCTGACGTTCGACGCTCTGTGAGATC 4777  
DB 4687 AGCTTTGTTCAGAGAGGATCCAGAGACGAAAAGCTGAGGTGACGCTCTGTGAGATC 4746  
QY 4778 GCGAAGTCCCGGACAGCAACGCTGCCAGAGCAACCAACCCATCTCGCTATGAG 4837  
DB 4747 GCGAAGTCCCGGACAGCAACGCTGCCAGAGCAACCAACCCATCTCGCTATGAG 4806  
QY 4838 CACCTGTTCGGGGGAGTGAATGGCGGTGACCTGTATGCAAGCAATCCCTGGGCTTTC 4897  
DB 4807 CACCTGTTCGGGGGAGTGAATGGCGGTGACCTGTATGCAAGCAATCCCTGGGCTTTC 4866  
QY 4898 GACCCCACTCCATACCCCGGAGCATCCCTCTGAGACGAGCGCTGCTACTACCTGCGCC 4957  
DB 4867 GACCCCACTCCATACCCCGGAGCATCCCTCTGAGACGAGCGCTGCTACTACCTGCGCC 4926  
QY 4958 GCAACCTTGGCCCCCAACCCCACTTACCCGACCTGTATCCAACCTTACTCATCCGCGCG 5017  
DB 4927 GCAACCTTGGCCCCCAACCCCACTTACCCGACCTGTATCCAACCTTACTCATCCGCGCG 4986  
QY 5018 TAACCCGACAGGGGGGCTGGAAGAACGGGACACATCATCAATGATCATCATCACTCG 5077  
DB 4987 TAACCCGACAGGGGGGCTGGAAGAACGGGACACATCATCAATGATCATCATCACTCG 5046  
QY 5078 CAGCAGATGCAACAACAAGGACCAACGCGCACTGAGCCCAAGAGCTGATGCTGAGGGGC 5137  
DB 5047 CAGCAGATGCAACAACAAGGACCAACGCGCACTGAGCCCAAGAGCTGATGCTGAGGGGC 5106  
QY 5138 CTCTGCGCCCGGAGTCTCGCTGGCACTCACTACGCTGGGGTCCCCAGAGCATATC 5197  
DB 5107 CTCTGCGCCCGGAGTCTCGCTGGCACTCACTACGCTGGGGTCCCCAGAGCATATC 5166  
QY 5198 GACCTGTCCCAAGTGCACAACCTGCTGTGCTCGTGGGCCCGGACACAGGGACCCCAAGC 5257  
DB 5167 GACCTGTCCCAAGTGCACAACCTGCTGTGCTCGTGGGCCCGGACACAGGGACCCCAAGC 5226  
QY 5258 ACCGCAATGAGACCGGCTTGTCTTACCTCCCAACCGGCGCCAGCCCTTGAAGAGCGGCAC 5317  
DB 5227 ACCGCAATGAGACCGGCTTGTCTTACCTCCCAACCGGCGCCAGCCCTTGAAGAGCGGCAC 5286  
QY 5318 AGCAGCTCCCACTCTCCCAAGAGTCCAAACAATTGAACAAAACAACAACGATCC 5377  
DB 5287 AGCAGCTCCCACTCTCCCAAGAGTCCAAACAATTGAACAAAACAACAACGATCC 5346  
QY 5378 TCGTCCGAGCGGGAGCGAGACCGGGATGAGAGGGGAGCCGGGATCGGGAAACGGGAAAG 5437  
DB 5347 TCGTCCGAGCGGGAGCGAGACCGGGATGAGAGGGGAGCCGGGATCGGGAAACGGGAAAG 5406  
QY 5438 TCCATCTCTCAAGTCCACAGACAGGTGAGACCGCACTTGAGAGACTTGTATACAGAG 5497  
DB 5407 TCCATCTCTCAAGTCCACAGACAGGTGAGACCGCACTTGAGAGACTTGTGTACAGAG 5466  
QY 5498 CAGAGCAGCGGACAGCGGAGCAAGCGCGGGGGTGGGGACAGACGCGCGCGCGC 5557  
DB 5467 CAGAGCAGCGGACAGCGGAGCAAGCGCGGGGGTGGGGACAGACGCGCGCGCGCGC 5526  
QY 5558 TCCCATCTCCCATGGCCACCAACATCTCGCCCATCTCCCTCGGACCCAGAGTCCCTTCAG 5617  
DB 5527 TCCCATCTCCCATGGCCACCAACATCTCGCCCATCTCCCTCGGACCCAGAGTCCCTTCAG 5586  
QY 5618 CAGAGACCCAATGTGCTTCAACAACAAGGATGAAGGATATCATCAACGCTGTGAGAGCC 5677  
DB 5587 CAGAGACCCAATGTGCTTCAACAACAAGGATGAAGGATATCATCAACGCTGTGAGAGCC 5646  
QY 5678 AGCAAGCCCAAGTCTGAGGTCAACTTCACCTTCACCCGTTTCGCCAGCTGCCACA 5737



Db 5647 AGCAAGCCCAAGCTGCTGAGGTCACTCCCTCAACCCGTTGCGCAAGCTGCACACA 5706  
Qy 5738 TTCCCACTGCGCAACCACTGCGCACTGCGCGGCACTCGAATGGGGCTTACCTTACCTC 5797  
Db 5707 TTCCCACTGCGCAACCACTGCGCACTGCGCGGCACTCGAATGGGGCTTACCTTACCTC 5766  
Qy 5798 ATGAGACCCGCTTGTGTGCGCAAGAGAGCCCGCGGTCGCGCGCAAGAGCCGCGCA 5857  
Db 5767 ATGAGACCCGCTTGTGTGCGCAAGAGAGCCCGCGGTCGCGCGCAAGAGCCGCGCA 5826  
Qy 5858 GCAAGACCGCGCACTGCTTCTTCTGCGCAAGCCCGCGCTTCGCGCGTGAAGCCGCGC 5917  
Db 5827 GCAAGACCGCGCACTGCTTCTTCTGCGCAAGCCCGCGCTTCGCGCGTGAAGCCGCGC 5886  
Qy 5918 TTCTCCCGCAAGAGGCTCGAGAGCCCGCGCTTGTGCTCTTCTTGTGCGCAAGCC 5977  
Db 5887 TTCTCCCGCAAGAGGCTCGAGAGCCCGCGCTTGTGCTCTTCTTGTGCGCAAGCC 5946  
Qy 5978 ACCATGCGCGCAACCCCTGCGAAGAACTTGCACCTTCAAGCGCAGCCCGCAAGCCGCG 6037  
Db 5947 ACCATGCGCGCAACCCCTGCGAAGAACTTGCACCTTCAAGCGCAGCCCGCAAGCCGCG 6006  
Qy 6038 GCGCCACTGCTCGGCTTGCAGCCCGCAAGCGGAAAGCTCAAGATTAACCTTTTTC 6097  
Db 6007 GCGCCACTGCTCGGCTTGCAGCCCGCAAGCGGAAAGCTCAAGATTAACCTTTTTC 6066  
Qy 6098 ATCCAGAACTGGAATCCGCTTCTGCGGTTACCAAGCGCAGCGCTTACAGCCCGCAAGG 6157  
Db 6067 ATCCAGAACTGGAATCCGCTTCTGCGGTTACCAAGCGCAGCGCTTACAGCCCGCAAGG 6126  
Qy 6158 GTGAGACCCGCTGAGCCCTGAGCTCAACCACTGAGCCCAAGAGGAGCTCCCAAG 6217  
Db 6127 GTGAGACCCGCTGAGCCCTGAGCTCAACCACTGAGCCCAAGAGGAGCTCCCAAG 6186  
Qy 6218 CACTTGAAGAGCTCGACAAAGAGCACTGAGAGGAGAGCTGCGGCCCAAGCGCAAGC 6277  
Db 6187 CACTTGAAGAGCTCGACAAAGAGCACTGAGAGGAGAGCTGCGGCCCAAGCGCAAGC 6246  
Qy 6278 CCCGTGAACCTTGGCGGAGAGCGGCCCACTCCCAACTGCGGCGCGCTGAGAGC 6337  
Db 6247 CCCGTGAACCTTGGCGGAGAGCGGCCCACTCCCAACTGCGGCGCGCTGAGAGC 6306  
Qy 6338 CAGCCCTGTCAGAGCCGCTGCTCAAGACCGGCCCAAGAGGTCAAGGTCAACGAGCGGT 6397  
Db 6307 CAGCCCTGTCAGAGCCGCTGCTCAAGACCGGCCCAAGAGGTCAAGGTCAACGAGCGGT 6366  
Qy 6398 GTCAACCTTGGCGGAGCACTGAGTATCAAGAGCTTCAAGAGCTTCAAGCGCAACCA 6457  
Db 6367 GTCAACCTTGGCGGAGCACTGAGTATCAAGAGCTTCAAGAGCTTCAAGCGCAACCA 6426  
Qy 6458 CAGAGACTCAGAGCAACCCCTGCGCGGCCCTTACTCTTCTTGGGAGCAGTGCCTCC 6517  
Db 6427 CAGAGACTCAGAGCAACCCCTGCGCGGCCCTTACTCTTCTTGGGAGCAGTGCCTCC 6486  
Qy 6518 GTCTGGAACCTCGCGGCCCAACCACTTACTCTTCTTGGGAGCAGTGCCTCC 6577  
Db 6487 GTCTGGAACCTCGCGGCCCAACCACTTACTCTTCTTGGGAGCAGTGCCTCC 6546  
Qy 6578 CCGGCGCGGTGCTCCCGCAAGCGAAGGAGGCAAGGCTTCAAGAGCAACCAAGAG 6637  
Db 6547 CCGGCGCGGTGCTCCCGCAAGCGAAGGAGGCAAGGCTTCAAGAGCAACCAAGAG 6606  
Qy 6638 TCGGTCTTGGGTGTGTGAGAGCGGTATTGAACCTGTGTCCCGAGCGGAGATGAG 6697  
Db 6607 TCGGTCTTGGGTGTGTGAGAGCGGTATTGAACCTGTGTCCCGAGCGGAGATGAG 6666  
Qy 6698 GAGCCAGGCACTCCCGAGTGTGTATCCCGCTGCTGATCCGGATGGGAAAGAGAG 6757  
Db 6667 GAGCCAGGCACTCCCGAGTGTGTATCCCGCTGCTGATCCGGATGGGAAAGAGAG 6726  
Qy 6758 GAGCCAGGAGATGGGCTCAAGTCTCAAGGCAACCAAGCGCGCAAGCTTCTTC 6817

Db 6727 GAGCCAGGAGATGGGCTCAAGTCTCAAGGCAACCAAGCGCGCAAGCTTCTTC 6786  
Qy 6818 AGCAAGCTGACGAGAGCACTCCGCAATGTGTAAGTCCAAAGAGAGATCAAG 6877  
Db 6787 AGCAAGCTGACGAGAGCACTCCGCAATGTGTAAGTCCAAAGAGAGATCAAG 6846  
Qy 6878 AAGCTGAACCCCAACCGGAAATGAGCTGAATTAATATCAAGCAAGCTGGAAGAG 6937  
Db 6847 AAGCTGAACCCCAACCGGAAATGAGCTGAATTAATATCAAGCAAGCTGGAAGAG 6906  
Qy 6938 ATCTTCAATATGCGCGCATCAAGGAAAGAGCTTATGACCTATAGAGCGCGGT 6997  
Db 6907 ATCTTCAATATGCGCGCATCAAGGAAAGAGCTTATGACCTATAGAGCGCGGT 6966  
Qy 6998 CAGGAATGCGAGCAACATGGGGCTGAGGAGCAATTAATGAAGGCACTATGGT 7057  
Db 6967 CAGGAATGCGAGCAACATGGGGCTGAGGAGCAATTAATGAAGGCACTATGGT 7026  
Qy 7058 AATATGACCAAGTGGAGAGTCCCGCGCTCAAGCGCAATGCTTTTAACTTGAAT 7117  
Db 7027 AATATGACCAAGTGGAGAGTCCCGCGCTCAAGCGCAATGCTTTTAACTTGAAT 7086  
Qy 7118 GCCAGTCCAGCTCCCGCTGCTATGCTTATGCTTATGCTGAGCGAGAGTACAC 7177  
Db 7087 GCCAGTCCAGCTCCCGCTGCTATGCTTATGCTTATGCTGAGCGAGAGTACAC 7146  
Qy 7178 ACACTACCTCGCAGGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7237  
Db 7147 ACACTACCTCGCAGGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7206  
Qy 7238 AAGCCAGTCCCGCGCGCGCGCGCTGAGCTGAGGAGAGGAGGAGGAGGAGGAGGAGGAG 7297  
Db 7207 AAGCCAGTCCCGCGCGCGCGCGCGCTGAGCTGAGGAGAGGAGGAGGAGGAGGAGGAGGAG 7266  
Qy 7298 GTGCACTGGAAGAGAGCTGCAACCGCGGAGCGCGCTCAACCAAGCGCTGTGGAGAGC 7357  
Db 7267 GTGCACTGGAAGAGAGCTGCAACCGCGGAGCGCGCTCAACCAAGCGCTGTGGAGAGC 7326  
Qy 7358 AAGCCCTGTCAGAGGTTCAAGGCAATCCCTTCAACCCCTGATCAAGCGGCTGAG 7417  
Db 7327 AAGCCCTGTCAGAGGTTCAAGGCAATCCCTTCAACCCCTGATCAAGCGGCTGAG 7386  
Qy 7418 GCGGTGTATGAGCTTCCCAACCCCAAGCGGCTTCCCGGAGGAGCGGCGCTTGT 7477  
Db 7387 GCGGTGTATGAGCTTCCCAACCCCAAGCGGCTTCCCGGAGGAGCGGCGCTTGT 7446  
Qy 7478 GCGCCCAACCAAGCTTGGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7537  
Db 7447 GCGCCCAACCAAGCTTGGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7506  
Qy 7538 CTCTCCGAGAGGAG 7552  
Db 7507 CTCTCCGAGAGGAG 7521

RESULT 10  
US-09-819-104A-4  
; Sequence 4, Application US/09819104A  
; Publication No. US20030027137A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, J. Don  
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES  
; TITLE OF INVENTION: AND USES THEREFOR  
; FILE REFERENCE: UMG-030  
; CURRENT APPLICATION NUMBER: US/09/819,104A  
; PRIOR FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: 60/193,138  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 8544  
; TYPE: DNA

ASN

ORGANISM: Mus musculus  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (160)..(7545)  
US-09-819-104A-4

Query Match 58.8%; Score 5037.2; DB 3; Length 8544;  
Best Local Similarity 78.3%; Pred. No. 0;  
Matches 6770; Conservative 0; Mismatches 1493; Indels 387; Gaps 45;

1 CATGTGCGGCTCCAGACAGAGCTTGGCAACAGACGAGGAGGACATGAGAGCCCGCTAACCC 60  
159 CATGTGAGATCCACACAGCCCTGTGGCAACAGATGGCGGCTGTGAGCCCGCTAACCC 218  
61 GCCCCACAGCCCTTCTCACTCCAGTCAGATCCCGGACGACACAGAGATGTGGGCTCT 120  
219 ACCCATGAGCATCTCTACCCCGGTCAAGATACCCGGTCCCAACAGAGATGGGCTGT 278  
121 GAGATACACAGCACTCTCCGCACTATGCTCCCACTTGTGCGGGCTCATCATCCA 180  
279 TGAGTACCAACACACCCCGGTACACACTCACACTGACCCGGTTCATCATCCA 338  
181 GCCCAGGCGGAGAGCCCTCCCTGTCTGTCTGATTCAGCCCGGAAATGAACGGTCCA 240  
339 GCCACAGAGAGGCGGCGCTCACTGTGTCAAGATTCAGCCTGGAGAGTGAACGGTCTCA 398  
241 GAGCTCCACCTGCGGACAGATCCCACTCATACCTGCGGAGCTGGGAAATCAGAGAT 300  
399 GAGGCTTCACTCGGCGCTGAGTCCCGACGTTCTGCTGTGAGCTGGGCAAGCCGACAT 458  
301 GAGATTCAATTGAAGCAAGCGCTCTGAGTGTGCTGTGACCCCTGTGCGGAC 360  
459 AGAATTCACCGAGACAAAGCGCCCGCTGAGACTACCCGATACCTCTGCGGCC 518  
361 GTACCCCTGTGTGGCAAGGCGCAAGCTCCGGAATCTGAAGACTCACCAAGAACCTTAG 420  
519 ATACCCCTGTGTGGCAAGGCGCAAGTGGGTGAAGACTTACCAAGAACCTTAG 578  
421 CCTGACGGGCAAGCTGGAACCGGTGTCTCCCGGAGCCCGGCACTGAACCTGAGCT 480  
579 CCTGCAAGGCAAGCTGAGGCTGTGTCACTCCAGTCCCGGCAAGCTGAGCTGAGCT 638  
481 GAGCTGTGCTCCGCAAGGCTGTGTCAAGAGAGAGCTGATCCAGAACAA--TGACCGCGT 537  
639 AGAGCTGCGCATCTGCACTGTCCAGAGAGAGCTGATCCAGAACAGATTGAGACGCGT 698  
538 GGAACGGAAGATCAACATGTGTAGACAGAGATCTCTAAGCTGAAGAAACACACACA 597  
699 GACCGTGAAGATCAACATGTGTAGAGCAGAGATCTTCAAGCTGAAGAAACACACACA 758  
598 GCTGAGAGAGAGGCTGCAAGCGCCGAGGCTGTGAAGACCGGTGTCAACGCGCCAT 657  
759 GTTGAAGAGAGAGGCGCCGAGGCTGCGGACCGAGAACCTGTGTGCGCACACCAT 818  
658 CGAGTGAAGACCGGCAAGCTGTGTGCAATCTTACGACGAGAACCGGAGAGAGGCTTGA 717  
819 AGAATCAAGACCGGAGGCTGTGTCAAGATCATCTTACGATGAGAACCGGAGAGAGCGGA 878  
718 AGCTGCAATCGGATTTGTGAAGGCTGTGGGCGCCAGAGGTGAGCTGCGCTGTACACA 777  
879 AGCCGCAACCGGATCTTGAAGGCTGTGGGCGCCAGAGGTGAGCTGCGCTGTACACA 938  
778 GCCCTCCGACACCGGAGATCATGAGAACATCAAAATTAACACGAGCATCGGAGAA 837  
939 GCCGCTGACACACCGCAGTACATGAATAATCAAAATTAACACGAGCATCGGAGAA 998  
838 GCTAATCTTTGTACTTCAAGAGAGAGATCAAGCTCGGAAACATGAGACAGAGATTCTG 897  
999 GCTAATCTTTGTACTTCAAGAGAGAGATCAAGCTCGGAAACATGAGACAGAGATTCTG 1058  
898 CCAAGGCTATGACCAAGCTCATGAGAGCTTGAAGAAAGAGTGAAGCGCATGAGAAACA 957  
1059 CCAAGGCTATGACCAAGCTCATGAGAGCTTGAAGAGAGTGAAGAGCGCATGAGAAACA 1118

958 CCGCGCGCGCGGCAAGAGAGCAAGTGCAGATTAACGAAGAAAGATTCTCTGA 1017  
1119 TCCGGAAGAGAGGCGCAAGAGAGCAAGTGTAGGAGTACTACGAAGAAAGATTCCCGGA 1178  
1018 GATCCGCAAGACGCGGAGCTGTGAGAGCCGATGAGAGCAGAGGTGGGCGAGCGGAG 1077  
1179 GATCCGCAAGAGACGCGGAGCTGTGAGAGCCGATGAGAGCAGAGGTGGGCGAGCGGAG 1238  
1078 TGGGCTGTCCATGTGTGCGCGCGCGGAGAGACAGAGGTGTCAAGATCATGATGCGCT 1137  
1239 TGGGCTGTCCATGTGTGCGCGCGCGGAGAGACAGAGGTGTCTGAGATCATGATGCGCT 1298  
1138 CTGAGAGAGAGAGAACTGTGAAGACAGATGCGGAGCTGAGCTCCGCGCCATCT 1197  
1299 GTGTGAGAGAGAGAACTGTGAAGACAGATGCGGAGCTGAGCTGAGCTCCGCGCCATCT 1356  
1198 GTACGAGCTGTACACGAGCGCATTAAGTTCATCAACATGAACGAGCTTTATGCGGACCC 1257  
1357 GTACGAGCGC-GACGAGAGAGATCCAGTTTCATCAACATGAATGATCTCATGATGAGCC 1415  
1258 CATGAAGGTGTACAAAGACCGCGAGGTATGTAACATGTGAAGTGAAGAGAGAGAC 1317  
1416 CATGAAGGTGTACAAAGACCGCGAGGTATGTAACATGTGAAGTGAAGAGAGAGAC 1475  
1318 CTTCCGGAAGAGTTTCATGACAGATCCCAAGAACTTTGGCTGATTCGATCATTCCTGGA 1377  
1476 CTTCCGGAAGAGTTTCATGACAGATCCCAAGAACTTTGGCTGATTCGATCATTCCTGGA 1535  
1378 GAGGAAGACAGTGTGTGAGTGTCTTATTTATTAATTAATTAATTAATTAATTAATTA 1437  
1536 GAGGAAGACAGTGTGTGAGTGTCTTATTTATTAATTAATTAATTAATTAATTAATTA 1595  
1438 TAAAGACCTGTGTGAGACGAGGTATGTGCGCGCGGAGAGAGAGAGAGAGAGAGAGAG 1497  
1596 CAGAGCTTGTGTGAGAGAGAGGTATGTGCGCGCGGAGAGAGAGAGAGAGAGAGAGAG 1643  
1498 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1557  
1644 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1703  
1558 GAAAGATGAAG 1617  
1704 GAAAGATGAAG 1763  
1618 GAAAGATGAAG 1677  
1764 GAAAGATGAAG 1823  
1678 CGAG 1737  
1824 TGAAG 1883  
1738 AGGCGCATTCACCGGCTTAATGTGTATGAGGCCCAACAGAGAGAGAGAGAGAGAGAG 1797  
1884 AGGCGGTATTCACCGGCTTAATGTGTATGAGGCCCAACAGAGAGAGAGAGAGAGAGAG 1943  
1798 GCAAGAGCGCGAGAGCTGTGCTCATGAGAGCTGAATGAGAGATTCTGCTGAGACAGAGAG 1857  
1944 GCAAGAGCTGTGAGAGCTGTGCTCATGAGAGCTGAATGAGAGATTCTGCTGAGACAGAGAG 2003  
1858 AATGGAACAGCAG 1917  
2004 GATGGAAGACAG 2063  
1918 GATGTTGGGCTTCAAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1977  
2064 CATGTTGGGCTTCAAG 2123  
1978 GCAAGAGCTGTGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2037  
2124 GCAAGAGCTGTGAGAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2183



Db 4189 CCACTCCCCCAACAACAACCTCGGGAAGCTGACTGAGACTTAAGCCCGGCCCTTGAC 4248  
Qy 4205 ---CTGGGCCCCCTGAAGCTGAAGCGGCGCCATGAAGGCTGTGTGGCAACGGTGAAGAG 4261  
Db 4249 CCTCTGGGCTCCCTGAAGCTGAAGCGGACTCAACGAGGGTGTGTGACCACTGTGAAGAG 4308  
Qy 4262 GGGGGCCGCTCCATCATGAGATCCCGCGAGAGAGCTGCGGCAACGCGCCGAGCTGGCC 4321  
Db 4309 GGGGCGCGCTATCATGAGATCCGAGAGAGAGCTGCGCGGCAACCTGAGCTACCC 4368  
Qy 4322 CTGGCCCCGCGCGCTCAAGAGGCTCATCAACGAGGCAACCCCGCTCAATGACAC 4381  
Db 4369 CTGGACAACAAGGCTCTGAAGAGGGTTCATCAACCAAGGCAACCCCACTCAAGTACGAC 4428  
Qy 4382 ACCGCGCGCTCCACCACTGGCTCCAAAAGCAAGAGTACGCTCCCTCATCGGCAAGCCCC 4441  
Db 4429 TCTGGGGGCACTCTCACTGGCAACAAGAAACAAGAGTGGCTCCATCATCGGCAAGCCCC 4488  
Qy 4442 GGGCGGAGCGTTCACCCCGTGCACCCGCTGAGTGTGATGGCCGACGCGCGGCACTGGA 4501  
Db 4489 GGC CGGCGCTTTCCTGCGCTCGACCCGCTGGAATATAGCTGAGCGCGGGGCACTGGAG 4548  
Qy 4502 CGTGCCTGCTACGAGAGAGGCTGAAGAGCGCGCCAGGGAACCGCCAGCACTGGGGGCG 4561  
Db 4549 CGTGCCTGCTATGAAGAGAGTCTGAAGAGCGGCTCAGGGAACGAGCTGTGTCAAGGGCG 4608  
Qy 4562 TCCATTGCGCGGCGCGCCCGGCTCATGTGCTGAGCTGGGTAAAGCGCGGCAAGCCCC 4621  
Db 4609 TCCATTCACAGTGGGGCTTCAGTCTGTGCTTGAACGTGGCAAGCCACGGAAGCCCA 4668  
Qy 4622 CTGACCTATGAGAGCAACGAGGCAACCTTTGCGGCACTCCCAAGAGTTTCGCGCTG 4681  
Db 4669 CTGACTTACGAAGACCAAGGGGCACTTTCACCAAGTCACTGCGACAGTGGCTCCCTGTG 4728  
Qy 4682 ACCATGCGGAGACCCACCGCGGCTTGACGAGAGGAGGCTTTGCTCCAGCAAGGATCC 4741  
Db 4729 ACCAAGAGGAGACCCACCGCAAGCTTCAGAAAGGAGGCTCTATCCAGCAAGGGGTCC 4788  
Qy 4742 CAGGACCGAAGACTGACGCTGCAAGCTGTGAGATGCGCAAGTCCCGGCAAGCAACCGTG 4801  
Db 4789 CAGGACCGGAGACTGACATCTACACCCCGGAGATGCGCAAGTCCCGCAACGACACTGTG 4848  
Qy 4802 CCGGAGCAACCAACCAACCTTCGCGCTATGAGCACTGCTTGCGGGCGTGAAGTGC 4861  
Db 4849 CCGGAGCAACCACTCAACCCCATCTCCCTATGAGCACTTGCTCGGGGCGTGAAGTGC 4908  
Qy 4862 GTGAGACTGTATGCAAGCAATCCCGCTGGGCTTGACCCCACTCCATACCCCGGCG 4921  
Db 4909 GTGAGACTGTATCGGTCAATCCCATTTGGCCTTTGACCCCACTCCATACCCCGAGG 4968  
Qy 4922 ATCCCTCTGGA---GCGAGCGGCTGCTACTACTGCGCGCGCAACCTTGCGCCCAACCCC 4978  
Db 4969 ATCCCTCTGGAAGCAAGACCGGAGCTTACTACTGCGCGCGCACTTGCGCCCAAGCCCC 5028  
Qy 4979 ACTTACCGGCACTGTACCAACCTTACTCATCCGCGCTACCCCGCAACGCGGCGCTG 5038  
Db 5029 ACTTACCAACCTGTACCACTTATCTCATCCGCGGCTTACCTGAGCAACGCGGCGCTG 5088  
Qy 5039 GAGAACCGGAGACCATCATCATATGACTCATCTCGACAGATGACCAACACGCT 5098  
Db 5089 GAGAACCGGAGACCATCATCATATGACTCATCTCGACAGATGACCAACACGCT 5148  
Qy 5099 GGCACCGGAGTGGCCGAGGAGCTGATGTGAGGGGGCTCGCGCGCGGAGTCTGCG 5158  
Db 5149 GCTTCGCGCATGGCCGAGGTGTGACATGTGAGGGGTCTGTCAACCGGAGAGTCTTG 5208  
Qy 5159 CTGGCACTCAACTGAGCTGGGCTCCCGAGGCACTCATGACCTGTCCCAAGTGCACAC 5218  
Db 5209 CTGGCGCTCAATTAATTCCGCTGGCCCAAGGCAATTAATGACCTGTGCCAAGTGCACAC 5268  
Qy 5219 CTGCTGTGTGTGTGCGCCCGACCAACGAGGCAACCCGAGCAACCGGCACTGAGCGCTTGTGC 5278  
Db 5269 CTGCGGTGTGTGTGCAACAGCGAGGCAACCCCTGCAACGCGCATTCGACGCGCTTGTGC 5328

Qy 5279 TACCTCCCCACCGCGCCCGACCCCTTTCAGAGCGGCAACAGAGCTCCCACTTCCCCA 5338  
Db 5329 TACCTCCCCACCGCGCCCGACCCCTTTCAGAGCGGCAACAGAGCTCCCACTTCCCCA 5388  
Qy 5339 GAGGCTCCACCACTTTCAGAAACCAACCAACAGCTCTGTCCGAGGGGAGGAGAC 5398  
Db 5389 GAGGCGCCCACTCACTTGTAAACCAACCTGCAATCTTCACTCGAGCGGAGCG--- 5444  
Qy 5399 CGGATTCAGAGCGGAGACCGGAGTGGGGAGGGGAAAGTCCATCTCACTGCAACAG 5458  
Db 5445 -----GAACTGAGCGGGAAGAGCAAGTCCATCTCACTGCACT 5490  
Qy 5459 ACGGTGAGCAACGCAACCACTTGAAGCTGTGAGAGTGAAGAGAGAGCGGCAAGCGCG 5518  
Db 5491 ACAGTGAGCAATGACCACTGAGAGCCTGTGAGAGAGAGAGAGAGCGGCGTGGGGCG 5550  
Qy 5519 AGCAGCGCGGGGGTGGGGGAGAGAGCGCGCGCGCTCCCACTCCATGCGCACAG 5578  
Db 5551 AGCAGCGG-----CGCGCTCCCAACCAACAG 5580  
Qy 5579 CACTGCGCATCTCCCTCGGAGCCGAGATGCGCTCAAGAGAGACCAAGTGTGCTTAC 5638  
Db 5581 CACTGCGCATCTCCCGGAGCCGAGAGCGCTTTCAGAGAGAGCGCAAGTGTGCTGAC 5640  
Qy 5639 AACACGAGCATGAAGGATATCATCAACCGCTGTGAGAGCCGAGCAACCCAGCTCTGAG 5698  
Db 5641 AACACGAGCATGAAGGAGTGTATCATCTCGTGAACCGGCAACCGTCTCGAGG 5700  
Qy 5699 TCCACTTCACCTCTCATCACCGCTTGGCCGAGCTGCAATTTCCACCTGCAACCACTGC 5758  
Db 5701 TCCACTTCACCTCTCGCTGTGCGCCAGCTGCAATTCACCTGCAACCACTGC 5760  
Qy 5759 CCACGTGGGAGCAACCTGATGGGGTCAACCTTACCTCATGAGAGCGGCTTGTGCGCC 5818  
Db 5761 CCACGTGGGAGCAACCTTGAAGGGGTCTACCTTACCTCATGAGAGCGGCTTGTACCC 5820  
Qy 5819 AAGAGGCGCCCGGGGTGCGCCGCGCAGAGCGGCCCGGAGAGACCGGCGCATGCTTC 5878  
Db 5821 AAGAGAGCTCTCGGGTGGCCGCGCGCAGAGCGGCCCGGTGTGAAGGTGGCATGCTTC 5880  
Qy 5879 CTGCGCAAGCCCGCAAGCGCGCTTCGAGGCGCGCTCTCCCGCAAGAGGCTGCG 5938  
Db 5881 CTGCGCAAGCCCGCGCGG-----GAGCGCGCTCTCTCAACCGCAAGAGGCTTC 5931  
Qy 5939 GAGCGCGCGCCCTGATGTGCTCTGTCTCTGCGCAAGCAACCATGCGCGCAACCCCTGCG 5998  
Db 5932 GAGCGCGCATCTTGAACCGCCCAAGCTTCAGGCAACAGCCATGCGCGCAACCCAGCA 5991  
Qy 5999 AAGAACTTCGCACTCAACCAAGCGCGAGCGCGCGCGCAACCTGCGCTGCGCTGCG 6058  
Db 6049 GATCTGACCGGAGAAAGACTCAAAAGTAAACCTTTTCATCAGAGAAATGGAACCTCGT 6108  
Qy 6119 TCTCTGGGTAAAC---GCGAGCACTTACAGCGCCCGAAGGGGTGGAAGCGGCTGACGCT 6175  
Db 6109 TCTCTGGGTAAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 6168  
Qy 6176 GTGAGCTCAACCAAGTGTGACCAACCAAGAGGGGCTCCCAAGCACTGGAAGAGCTGAC 6235  
Db 6169 GTGAGCTCCCGCAAGCTGACCAACCAAGAGGGGCTCTCCAAACCTTGGAAAGCTGAG 6228  
Qy 6229 AAGAGCACTTGAAGAGGAGGTGCGGCAAGAGCAAGCGCGCCCATGAAAGCTGACGCG 6288  
Db 6236 AAGAGCACTTGAAGAGGAGGTGCGGCAAGAGCAAGCGCGCGCTGAAAGCTTGGCGG 6295  
Qy 6236 GAGGCGCGCACTTCCCAACCTGCGGCGCTGCTGAGAGGCAACCGCTGCTGACGCGG 6355  
Db 6289 GAGGCTGCGCATCTCCCAATCTGCGGCGCACTGCGGAGAGCAACCGCTCATTCAGCCCA 6348

6356 CTGCTCAGACCGCCCGAGGGGTCAAAAGTCAACAGCGGGTGTCACTCTGCCCGACAC 6415  
6349 CTCTCCAGAGCTGCCCGAGGCATCAAAAGTCAACAGAGGGTGTCACTCTGCCCGAC 6408  
6416 ATCAGTGAAGTATCAACAGAGTCAACCGGCAACCCACAGAGCTCAGCGCACCC 6475  
6409 ATCAGGAGGTCAATTACGAGAGTCAACCCGGCACCAACCGAGAGCTCACTGGCCCC 6468  
6476 CTGCGCGCGCGCTCACTCTCTCCGCGGGGAGAGCGCCCGTCCGAGCTCCGCGCG 6535  
6469 CTTCGCGCGCTCTCACTCTCTCTCCGAGCGAGCTGCGCTGTGTGATCTTCGCGCG 6528  
6536 CCAACCGAGTGAACCTTCACTCTCCGCGCGAGCATGTGTGCGCGCGCTGCGCTCCCC 6595  
6529 CCAACCGAGTGAACCTTCACTCTCCGCGCGAGCATGTGTGCGCGCGCTGCGCGCTCCCC 6588  
6596 CACAGCGAAAGGGGGCAAGAGTCTCAAGAGCCAAACAGAGCTGTGGTCTTGGGTGTGT 6655  
6589 CACAGTGAAGGGGGCAAAAGGTCCCGAGAACCGAGCAAAACATCGGTCTGGGCAAGAT 6648  
6656 GAGGAGGTATTGAACCTGTGTCCCAACCGAGAGGGATGAAGGAGCGAGGCACTCCCG 6715  
6649 GAGGAGTCAATTGAGCGCTGTGTCCCAACAGAGGGATGACTGAGCGAGGATGCTCGG 6708  
6716 AGTCTGTGTACCCGCTGTGTACCGGGAATGGGAAACAGAGCGCCAGAGGATGGG 6775  
6709 AGGCTGTGTACCCAGCTGTGTATCGAGAGGGGAACAGGGCGAGCC--CAGGATGGG 6765  
6776 TCCAGTCTTCAGGCAACACAGCGAGCGCGCAAGCTTCTTCAGCAAGCTGACGAGG 6835  
6766 TCTAAGTCTTCAGGCAACACAGCGAGCGCGCAAGCTTCTTCAGTAAAGCTGAGAGG 6825  
6836 AACTCCGCAATGTGTCAAGTCAAGAGAGAGATCAACAGAAAGCTGAAACACCCCAAC 6895  
6826 AACTCCGCAATGTGTGAAGTGAAGAGAGAGATCAACAGAAAGCTCAACACCCCAAC 6885  
6896 CGGAAATGAGCTGTAATCAATATCAGCAGCGCTGGAGCGAGATCTTCAATATAGCCGCG 6955  
6886 CGGAAATGAGCTGTAATCAATATTTGGCGAGCTGGAGCGGAAATCTTCAACATAGCCGCG 6945  
6956 ATTCACCGGAACAGGCTTATGACTATAGAGCCAGCGGCTGAGGAACATGCGAGCACC 7015  
6946 ATTCATGAGAGAGGCTTATGACTGTAGAGCCAGCGGCTGAGGAACATGCGAGCACC 7005  
7016 AACTAGGGGCTGAGAGGCAATTAAGAAAGGCACTATGGGTAAATATGACAGTGGGAA 7075  
7006 AACTAGGGGCTGAGAGGCAATTAAGAAAGGCACTATGGGTAAATATGATGATGGGAA 7065  
7076 GAGTCCCGCGCGCTCAGCGCAATGCTTTAACTCTGAAATGCGAGTGCAGCTGCC 7135  
7066 GAGTCCCGCGCGCTCAGCGCAATGCTTTAACTCTGAAATGCGAGTGCAGCTGCC 7125  
7136 ---GCTGCTATGCGCAATTAACCGCTGTGACGAGCGAGTGAACCACTACCTGCCCA 7192  
7126 GCGTGTGCTATGCGCAATTAACCTGTGACGAGCGAGTGAACCACTACCTGCCCA 7185  
7193 GGTGTGCGGGGAAGGCAAGGTCTCTGGAGACCCAGCAGCGGAAAGCAAGTCCCG 7252  
7186 GGTGTGCGGGGAAGGCAAGGTCTCTGGAGACCCAGCAGCGGAAAGCAAGTCCCG 7245  
7253 GCGCGCGGCGCTGCACTGCGGAGCGGAGCAACCTGTGTCTCTCAGTGCATGAGGGA 7312  
7246 GCAACGAGGCTTACGCTCGAGAGCCAGCCCTTCTGTCTCTCAATGATCACTAGAGGG 7305  
7313 GACTGCAACCGCGGAGCGCGCTCACCAACCGGCTGTGGAGAGAGAGCGCTGTCCGCA 7372  
7306 GACTGCAATCGCGGAGCACCACTCAACCAACGCTGTGTGGAGAGAGCGCTGTCCGCA 7365  
7373 GGTTCACAGCGCATTTCCCTTCAACCCCTGTATATATGAGGCTGAGAGCGGTGTCAATG 7432  
7366 GGTGTACAGCGCATTTCCCTTCAACCCCTGTATATATGAGGCTGATGTCATGAGCC 7425  
7433 TCCCAACCCCAACGAGGCTCTCCCGGAGGAGCGGCGCTCGCTGAGCCCAACGAGCC 7492

7426 TCCCGCGCGCGCACTGAGCTTGTGCGAGGAGCGGCGCCCTTACTGTGCTCCCAACGAGCC 7485  
7493 TGGAGAGAGAGCGCCAAAGCACTGTCTGTGCAATTAAGAGCACTCTCCGACAGGAG 7552  
7486 TGGAGATGAAGAGCCAAAGCACTGTGTGTCAAGATTAAGACACTCTGAGACAGGAG 7545  
7553 TGAATCAGAAAGAGCGGGGGGGGGCGGGCGGTGEBURPERLUTGAGGTCCGAGGAGC 7612  
7546 TGAACAGGATTTGGGGGGAGAGCGG-----TGCAAGGTCCCGACAA 7586  
7613 CACAGGAACGCGCTTGCAGAGCGGGGGCGCTGCCAATCTCCCAACAAAGAAAGAGCC 7672  
7587 GCGAGAAAGAGCGCCAGACATGAGAGAGAGAGAGAGCTGTGAGCTCGGAGACTGAGAAAGAGCC 7646  
7673 CCTGAATCCGCGCTGC-----GCTTCATCATCTGTCCGTCCGAGAGCGG 7717  
7647 CTGAAGTGTGCGCTGTGCTGATCCGTCCGTCCGTCACTATCTGTCCATTCAGAGCTGG 7706  
7718 CATCTTGCCTGTCTTAAAGCTTAACTTAAAGATCTCCGCGCGGGCTGCGCTGTGAG-- 7775  
7707 CATCC--TGCTGTCTTAAAGCTTAACTTAAAGATCTCCACCCCGGGCTGCGCTGCGAGTG 7765  
7776 --ACCTTACTCAGGGAGTGTTAACCTGTGTCTCGGAAAGGAGGAGAA----- 7821  
7766 ACCTTACATCAGGGAGTGTTAACCTGTGTCTGAGAGGGGAGATGAGACAGAGGGAGG 7825  
7822 ----GGGCGCGGAGAGGGGGGAGCGGAGCGGTGTGGAGCCACACAGCGGGCCAGGG 7876  
7826 ACAGCGGGGCGAGAGGGGGGGGAGCAGCAATCGTGTCAAGTGCACATCGGTCACTGGGG 7885  
7877 CGGCGAGGAGCCCAAGAGAGATGACAGCAGCACTTCAGCGCCAGCTGCT--CCCGCAAT 7934  
7886 TACGCGGGAGCCCAACAGAGCTGACAGGCACTTCATGCAACGCGCTCGCCCTTACC 7945  
7935 GCATTTGAAACCAAGCTTAACTGAGCTGCGAGCCCGCGCGCTCTCGCGCTTCCA 7994  
7946 GCATTTGAAACCAAGCTTAACTGAGCTGCGGAGGTCTCGCGCTCTCTCTGCCCCCA 8005  
7995 TCCCGCTTACGCGCTGTGAGCAGATGAGCGAGCCCTGTTCAGAGCCCGAGTGGCTCGTT 8054  
8006 GCGCGCTTACCTGTGAGCAGAGAGAGCTTCCAGCTTATCTGCGCTTAACTGCTGATCA 8065  
8055 CGGCTCCCAACAGATGCGCCAGCGCAAGAGATGCTGGAACCAAGTCAAGCGCAGGTGG 8114  
8066 TCGAGATCTCCAA--GGCACCAGCCCAAGAGCTGGAGCCCATGAGACAGAGTGG 8122  
8115 GCGGACAAAGGCGCAGGTGCGGCTGAGGGGAAACGAGATGCTCCAGAGACTGAGCTGTTT 8174  
8123 GTGACACAAAGGGGCGTGG-----CTGGGGCAGGATGCTTGAGGAACTGAGCCGTT 8174  
8175 TTTTACACATGCTTGCAGAGCGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8234  
8175 TCCCGGCTGTGTGCTGTGCAACGCG---GAGGGAAGGAGCAGGTGAAATG--GTGTTGCG 8228  
8235 TTAAGGATTAATTTTGAATTAACCTTCAATGAATTAATTCAGATGTTTTCGAAAGGAGG 8294  
8229 TTAAGGATTAATTTTGAATTAACCTTCAATGAATTAATTCAGAGTCTCAACAGAGAGG 8288  
8295 ACTTACCAAGTATTAATGCTGTGTGCTTTTGAATCTCTGCTTAACTGTTCAAGAGGAGTGT 8354  
8289 ACTCGGCC--GTGTTCTCCGCGTGTGC--TTTGTGTCTAAGCTATGTTTCAAGAGGACGT 8346  
8355 GCAAGCGCAAGTGGTGAACCCCATATCTGCGAGACCAAGGGGGCGGGAGCTGTCTGTC 8414  
8347 GC---CAGCCAAAGCGGTGGCCCACTATGCGAGGACTTGTGGGGGTCAAGGGGCTCCGAGAC 8403  
8415 AGGCGCGCGGTGTCTCTCCCTCCCTCCCTTCTTGGGAGGAATGAATTAATGATGATGATTC 8474  
8404 ACAGCACTGTG--CCCTTCCCAACCTTATCTTCAAGCAATGAGCTGATGATGATTC 8460  
8475 TGTGCGCGCATTTGTGCAAGGAGGTGTGATTTCTGTCAATTAACAGCTGTTCTAATTA 8534

Db 8461 TGTCGCCCTCTCTGTGCA--CGCGGCANTTGTCTATTACATGTTGTTCCAAATTA 8517  
Qy 8535 AAAAGCGAAT 8544  
Db 8518 AAAAGCAAAAT 8527

RESULT 11  
US-10-764-420-1526  
; Sequence 1526, Application US/10764420  
; Publication No. US20050084872A1  
; GENERAL INFORMATION:  
; APPLICANT: Lum, Pek Yee  
; APPLICANT: Tan, Yejun  
; APPLICANT: Dai, Hongyue  
; TITLE OF INVENTION: Methods For Determining Whether An Agent  
; FILE REFERENCE: ROSA122057  
; CURRENT APPLICATION NUMBER: US/10/764,420  
; PRIOR FILING DATE: 2004-01-23  
; PRIOR APPLICATION NUMBER: US 60/442,797  
; PRIOR FILING DATE: 2003-01-24  
; PRIOR APPLICATION NUMBER: US 60/474,413  
; NUMBER OF SEQ ID NOS: 3683  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1526  
; LENGTH: 8544  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-764-420-1526

Query Match 58.8%; Score 5037.2; DB 9; Length 8544;  
Best Local Similarity 78.3%; Pred. No. 0;  
Matches 6770; Conservative 0; Mismatches 1493; Indels 387; Gaps 45;

Qy 1 CATGTGCGGCTTCAACAGCTTTGTGGCACAGCTGAGAGGCGCACTGAGCCCGCTACCC 60  
Db 159 CATGTGCGGCTTCAACAGCTTTGTGGCACAGCTGAGAGGCGCACTGAGCCCGCTACCC 218  
Qy 61 GCCCGACAGCCCTTCTACCGAGTGCAGTGGCCGCGACGACAGAGGCTGGGCTCCT 120  
Db 219 ACCCATGACATCTCTACCGAGTGCAGTGGCCGCGACGACAGAGGCTGGGCTCCT 278  
Qy 121 GGAGTACAGACACACCTCCGAGTATGCTCCACCTGTGCGCGGCTCCATCATCA 180  
Db 279 TGAGTACAGACACACCTCCGAGTATGCTCCACCTGTGCGCGGCTCCATCATCA 338  
Qy 181 GCCCGACAGCGCGAGGCTCTCTGTGTGTGAGTTCCAGCCCGGAAATGAACGGTCCA 240  
Db 339 GCCCGACAGCGCGAGGCTCTCTGTGTGTGAGTTCCAGCCCGGAAATGAACGGTCCA 398  
Qy 241 GGAGTCAACCTCGGGCGAGTCCCATCTACCTGCGCGAGCTGGGAAATGCAGAGT 300  
Db 399 GGAGTCAACCTCGGGCGAGTCCCATCTACCTGCGCGAGCTGGGAAATGCAGAGT 458  
Qy 301 GGAGTTCAATTGAACAGAGCGCTCTGAGTGTGCTGACCCCTGCTGCGACC 360  
Db 459 AGAATTCAACGAGACAGAGCGCTCTGAGTGTGCTGACCCCTGCTGCGACC 518  
Qy 361 GTCACTGCTGTGCGACAGCGGACAGCTGCGGAGTGTGAAGCTTCAACAGAGCCGTAG 420  
Db 519 ATCACTGCTGTGCGACAGCGGACAGCTGCGGAGTGTGAAGCTTCAACAGAGCCGTAG 578  
Qy 421 CCTGACGGGCAAGCTGGAACGGGTGTCTCCCGACGCCCCCGACACTGACCTGAGCT 480  
Db 579 CCTGACGGGCAAGCTGGAACGGGTGTCTCCCGACGCCCCCGACACTGAGCT 638  
Qy 481 GGAGCTGTGCGCGACAGCTGTCCAAAGAGAGCTGATCCAGAAACA--TGAACCGGCT 537  
Db 639 AGAGTGTGCGCGACATCTCACTGTCCAGAGAGAGCTGATCCAGAAACAAGTTGACCGGCT 698  
Qy 538 GGACCGAGAGATCAACATGTGTAGAGACAGAGATCTCTAAAGTGAAGAAAGACAGCAACA 597

Db 699 GACCGTGAAGATCAACATGTGTAGAGACAGCAAGATCTCCAAAGCTGAAGAAAGAGCAACA 758  
Qy 598 GCTGGAGAGAGAGCTGCGCAAGCGCCGAGCTGAGAAAGCCGTGTGACCGCGCCCAT 657  
Db 759 GTTGGAGAGAGAGCGCGCAAGCGCCCGCAACCGAGAAAGCTGTGTGCGCACACCCAT 818  
Qy 658 CGAGTGAAGACACCGACCTGTGTGACATCATCTTACACAGAAACCGGAAGAAAGCTGTA 717  
Db 819 AGAATCAAAAGACCGAAGCCGTGTGACATCATCTTACATGAAGAACCGGAAGAAAGCCGA 878  
Qy 718 AGCTGCAATCGAGATTCGTGAAAGCGCTGGGCGCCAGGTGAGAGCTGCGCTGTACACA 777  
Db 879 AGCGCACACCGAGATCTTGAAGGCTGGGCGCCAGGTGAGAGCTGCTGTGTACACA 938  
Qy 778 GCCCTCCGACACCGCGAGTATCATGAGAAATCAAAATTAACAGCGATGCGGAAGAA 837  
Db 939 GCCGTGACACACCGCGAGTACATGAAATCAAAATTAACAGCGAGTATGCGGAAGAA 998  
Qy 838 GCTAACTTGTGATCTTCAAGAGAGAGATCAAGCTTGGAAACAATGAAACAGAGTTCTG 897  
Db 999 GCTGATCTTGTGATCTTAAAGCGAGAGAACAGCGCGCAAGCAGTGGGAACAGCGCTTCTG 1058  
Qy 898 CCAAGCGTATGACCAAGCTCATGAGGCGCTTGGAAAAAAGGTGAGAGGCACTGAAAAA 957  
Db 1059 CCAAGCGTATGACCAAGCTCATGAGGCGCTTGGAAAAAAGGTGAGAGGCACTGAAAAA 1118  
Qy 958 CCGCGCGCGCGCGCAAGAGAGAGAGAGAGAGTACAGAAAGAGTTCCTGTA 1017  
Db 1119 TCCCGGAAGAGAGGCGCAAGAGAGAGAGAGTACAGAAAGAGTTCCTGTA 1178  
Qy 1018 GATCCGCAAGACGCGCGAGCTGTGAGAGCGCATGTGACAGCAGGTGGCCAGCGGGCGAG 1077  
Db 1179 GATCCGCAAGAGCGCGAGCTGTGAGAGCGCATGTGACAGCAGGTGGCCAGCGGGCGAG 1238  
Qy 1078 TGCGGTTCATGTCGCGCGCGCGAGAGACAGAGGTTCAGAGATCATGATGAGCT 1137  
Db 1239 TGCGGTTCATGTCGCGCGCGAGAGAGAGAGTTCAGAGATCATGATGAGCT 1298  
Qy 1138 CTGAGAGCAGAGAACTGTGAGAAAGCAGATGCGCAGCTGCGCGTATCCCGCCCATCT 1197  
Db 1299 GTCTGAGAGAGAACTGTGAGAAAGCAGATGCGCAGCTGCGCGTATCCCGCCCATCT 1356  
Qy 1198 GTACGAGCTGACAGACGCGCATCAAGTTCAATCAATGAACGGGCTTATGCGCGAC 1257  
Db 1357 GTACGAGCGC-GACGAGCAGAGATCAAGTTCAATCAATGAATGAGATGAGAC 1415  
Qy 1258 CATGAAGGTGTAACAAGACCGCCAGGTCAATGAACAATGAGTGAAGAGAGAGAC 1317  
Db 1416 CATGAAGGTGTAACAAGACCGCCAGGTTCACAACATGTGAGCGCAGAGAGAGAGAC 1475  
Qy 1318 CTTCGCGGAGAAAGTTCAATGACAGCATCCCAAGAACTTTGGCTGATCGCATCTCTGGA 1377  
Db 1476 CTTCGCGTGAAGAAAGTTATGACAGCACCTTAAGAACTTTGGCTGATCTCTGGA 1535  
Qy 1378 GAGGAAGACAGTGTGAGTGTCTCTATTTACTTACTTGAAGAAATGAGAACTTA 1437  
Db 1536 GAGGAAGACAGTGTGAGTGTCTCTATTTACTTACTTGAAGAAATGAGAAATTA 1595  
Qy 1438 TAAAGAGCTGTGAGAGCGGAGCTATCGGCGCGCGGCAAGAGCGACGAAACAACAGA 1497  
Db 1596 CAAGAGCTTGTGAGAGCGGAGCTATCGGCGCGCGTGAAGAGC-----CAGCA 1643  
Qy 1498 GCAGCAGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1557  
Db 1644 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1703  
Qy 1558 GAAAGATGAAG 1617  
Db 1704 GAAAG 1763  
Qy 1618 GAAAGCAG 1677



Db 1764 GAAAGAAAGAAAGAACTCAGCAAGAGAAAGACAGACGACCTTCTGGCGAGGACAAACGA 1823  
 Qy 1678 CGAAGAGAGGCTGTGGCTCCAAAGGCGCAAACTGCGCAACAGCGAGGAGAGCGCA 1737  
 Db 1824 TGAAGAAAGAGGCGGTGGCTCCAAAGGCGCAAACTGCGCAACAGCGAGGAGAGCGCA 1883  
 Qy 1738 AGGCGGCATCAAGCGCTCAATGGCTAATGAGGCAACAGGAGAGGCGCATCAAGCGCGCA 1797  
 Db 1884 AGGCGGTATCAAGCGCTCCATGGCGCAACAGGCGCAACATGAGAGAGACAGCGCGCA 1943  
 Qy 1798 GCAAGAGCGCGAGCTGGCTCCATGAGAGCTGAATGAGAGTTCCTGCTGAGCAGAAAGAA 1857  
 Db 1944 GCMAATTCAGAGCTGGCTTCATGAGAGTGAACGAGAGTTCCTGCTGAGCTGAAGAAAG 2003  
 Qy 1858 AATGGAACAGCGCAAGAAAGTCTCCGAGAAACAGCGCGCACTGGTCGGCGCATGCGCG 1917  
 Db 2004 GATGAGACAGCAAGAAAGGCTCTTGAAACATGGAGAGAACTGGTCAGCGCATTTGCGCG 2063  
 Qy 1918 GATGTGGGCTCCAGACTGTGTGCGAGTGTAAAGAACTTCACTTCACTTAAGAAAG 1977  
 Db 2064 CATGTGGGCTCCAGACCGTGTCCAGGTAAAGAACTTCACTTCACTTAAGAAAG 2123  
 Qy 1978 GCAAGAACTTCAGATGATCTTGCAGACAGCAAGCTGAAAGATGAGAGAGAGAAACG 2037  
 Db 2124 GCAGAACTTGAACGAAATCTTCAGACAGCAAGCTTAAAGATGAGAGAGAGAAACG 2183  
 Qy 2038 GGGAGGAAGAAAGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2097  
 Db 2184 TCGAGAGAAAGAAAGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2243  
 Qy 2098 GGTGAGAGATGAGAGATGAGAGCGCTCGGCGCTGAGCGGAAATGAGAGAGAGATGAG 2157  
 Db 2244 CGGTGAGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 2303  
 Qy 2158 GAGAGCTGAAGCTTATCATGCTCTTGGAAATGAGAGTGGCGCAAG--GGAGATGCACTG 2214  
 Db 2304 GAGAGGAGAAAGCGCTCAAGGCGCTCTGGAAATGAGAGTGGCGCAAGTGGGAGTGCAGT 2363  
 Qy 2215 CCGAGGCACTTCAAAAGAGCTCAGACCGGAGAGATCCCTCTCTCAACATGAGG 2274  
 Db 2364 CCGAGTGTCTCAAAAGAGCTCTGATCTGAGAGTGTCCATCCCGGCTTCAAGAG 2423  
 Qy 2275 CCGCAAGACACAGGCGAGATGGGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 2334  
 Db 2424 CAGGAAGAGAC-----TGGGCTTAAACCATCTGGGCACTGAAGCATTTGCCCGCTG 2474  
 Qy 2335 ACCCGGAGGCG 2394  
 Db 2475 CACCGGCGCGCGCTTCTCTCTCAAGAAAGACCGGCGCGCGCGCGCGCGCGCGCG 2534  
 Qy 2395 GGCCTTGAAGCGCGAGGCGCGCTTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2454  
 Db 2535 AGTCCCTGATGCGCA---GTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2591  
 Qy 2455 TCCTCTGTGTGCTCCCAAG 2514  
 Db 2592 CCGGCTTATCTGTGAACAGAGATGAACAGAGCGCGCGCGCGCGCGCGCGCGCGCG 2651  
 Qy 2515 GGAAGGAGAGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2556  
 Db 2652 AGATGCGCAAGAGAGAGATCTGAGGCGCGAGAGATGATGTGGAAGAGCGAGAGAGCG 2711  
 Qy 2557 AGTGAACAGAGAGAGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2616  
 Db 2712 CAGAGGCTTGAAGAGCGCGCGAGAGTGAAGAGTGAACAGAGAGAGAGAGCGAGAG 2771  
 Qy 2617 GGGGCGCGCG---CAAGGCGAGAGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 2673  
 Db 2772 AGAGCTTGAAGAGAGAGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2831  
 Qy 2674 CAGAGGAGAGAGAGAGAGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2733  
 Db 2832 TAAAG---GTGAGAGAGGCTGTGTGAGCAAGGCGCTGTGAGCAAGGCTTCCAGCTCAG 2888

Qy 2734 CCGCGGAG 2793  
 Db 2889 CACCGGAG 2948  
 Qy 2794 CCGGAG 2853  
 Db 2949 AGGTGAAG 3008  
 Qy 2854 CCGGCGAG 2913  
 Db 3009 CCGGCGAG 3068  
 Qy 2914 TGCATATCCCGCGAG 2973  
 Db 3069 CGGCGAG 3125  
 Qy 2974 CACCAAG 3033  
 Db 3126 CCGAAAGCGAG 3185  
 Qy 3034 CCGTGAAG 3093  
 Db 3186 GTCTCAG 3245  
 Qy 3094 CAGAGAGC-----CTTGCAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3130  
 Db 3246 GAAAG 3305  
 Qy 3131 GAGGAG 3190  
 Db 3306 ACTGAG 3360  
 Qy 3191 AAGGCTCCCGGAG 3250  
 Db 3361 AAGACTTTCAG 3414  
 Qy 3251 CCGGCGAG 3310  
 Db 3415 CTGCTTGGGCTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3471  
 Qy 3311 AACCGGCTCCCGTATCTCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3370  
 Db 3472 AACCGGCGAG 3531  
 Qy 3371 GCGATTTCCAG 3430  
 Db 3532 GCGATTTCCAG 3588  
 Qy 3431 CCGGAG 3490  
 Db 3589 CCGATGGGCGCTTCAAG 3642  
 Qy 3491 TTCAAG 3550  
 Db 3643 ----- 3642  
 Qy 3551 CTGGGAG 3610  
 Db 3643 -----GGAG 3662  
 Qy 3611 CCGGCGAG 3670  
 Db 3663 ACCAGTGAAG 3715  
 Qy 3671 ACATACCGGAG 3730  
 Db 3716 AGGTGAAG 3774  
 Qy 3731 ACCAGATCATGGAG 3790  
 Db 3775 AGCAGATGATGGAG 3834

OY	3791	CCCAAGGGCCACGTATCTTAGAAGGACAAAGAGGGCCACGTCCTTGTCTTATGAGGGGTGGC	3853
Dp	3835	CCCAAGGGCCACGTATCTTAGAAGGGCAAGAAAGGCCACGTCCTATCTTAAGAAAGTGTGT	3894
OY	3851	ATGTCGTGACCCAGTGTCTCCAAAGGAGACGGCACAAGCAGCTCGAGACCCCCCATAG	3910
Dp	3895	ATGTCGTGTACAGTGTCTTAAGAGGATGGAAGAGACGTCGGGGCCACCCCATAG	3955
OY	3911	ACGGCCGGCCCCAAGCCGACCTATGACATGATGAGGGGCCGGTGGGAGAGCCATCTCC	3970
Dp	3955	ACTGCGGCCCTTAACGCACCTATGACATGATGAGGGGCCGGTGTAGGAGACTGTCAAC	4014
OY	3971	TCAGCCACGATCGAAGGTCTCATAGGGCCGTGTGCATCTCCGCCGAGACCAAGCCCCAC	4033
Dp	4015	TCAGCCACGATCGAAGGACTCATAGGGCCGGCCGACATCCCTGAGCAGC-----ACAGCCCC	4068
OY	4031	CACCTCAAAAGACAGACCAACATCCGGGGGATCATACACAAGAGGATCCCTCGGCTCAC	4099
Dp	4069	CACCTCAAAAGAGACGATCATATCCGAGGCTTCATCAGCAAGGACATCCGAGGTCCTAT	4128
OY	4091	GTGAGAGCAGAGAGACTACCTGTGGTCGGAGAGCCAAAGCTCTAAACGGGAGGGACAG	4156
Dp	4129	GTGAGAGCGCAGAGAGACTATCTTAAGCGGGAGGCCAAAGCTCTTAAGCAGAAAGGACA	4186
OY	4151	CCTCGGCCCCCACCGCCCTCATCGGGACTGTACCGAGGCTTACAGACGCAAGCC-----	4204
Dp	4189	CCACTCTCCCCACCAACCACTCTGGGACCTGTACTGTAGACTTACAAAGCCCGCCCTGTGAC	4248
OY	4205	---CTGGGCCCCCTGAAGCTGAAGCCGGGCCCATGAGGGCCGTGTGGCCACGGTGAAGAG	4261
Dp	4249	CCTCTGGGATCCCTTAAGCTGAAGCCGACTCACGAGGGTGTGTAGCACTGTGAAGAG	4308
OY	4262	GGGGCCGCTCCATCATGAGATCCCGCGAGGAGCTGTGGCACAACGCCGAGCTGTGCC	4321
Dp	4309	GGGGCCGCTGTATCATGAGATCCCGAAGAGAGGCTGTGGCCGACACTGAGGTTACCC	4366
OY	4322	CTGGCCCCCGGCCCGCTCAAGAGGGCTTCATCAGCAGAGGCACTCCGCTCAATGACAC	4381
Dp	4369	CTGGACACACGGCCCTCTGAAGAGGGTTCATCAACCCAGGGCACCCCACTCAAGTACAC	4428
OY	4382	ACCGGCGGTCACCACTGGGCTCCAAAAGACAGAGTACGTCCTCATGGGACGGCC	4441
Dp	4429	TCCTGGGGCACTCTCACTGTGGACCAAGAAACAGCAGTGGCTCATATGTGGACGGCCC	4488
OY	4442	GGCCGGAGCTTCCACCCCGTGCACCCGCTGATGTGATGTGCGCAGCGCCCGGCACTGANA	4501
Dp	4489	GGCCGGGCTTTTCCGTGCCCTGACCCGCTGGACCAATATGTGTGACGCCCGGGCACTGGAG	4548
OY	4502	CGTGCCTGTCTAGAGAGAGCTTGAAGAGCCGGTCAAGGACACGACGATGTGCGAGGGCC	4561
Dp	4549	CGTGCCTGTCTAGAGAGAGCTTGAAGAGCCGGTCAAGGACACGACGATGTGCGAGGGGCC	4608
OY	4562	TCGATTTGGCCCGCGGCCCGGTCAATTGTGCTGAGCTGGGTAAAGCCCGGCAAGACCCC	4621
Dp	4609	TCGATTCACGTTGGGGCTTCAGTGTGTGTGCTTGAATGTGGGAAACCAACGCAAAAGCCCA	4668
OY	4622	CTGACTTATGAGACCAAGGGGCACTTTTGCCGGCCACTTCCACGAGGTTTGGCCGTG	4681
Dp	4669	CTGACTTATGAGAGACCAAGGGGCACTTTCACAGTACCTGTGCCACGTGGCTCCCTGTG	4728
OY	4682	ACCATGCGGAGAGCCACGCGCGGCTGACAGAGGGCAGCTTTTGTGTCCAGCAAGCATCC	4741
Dp	4729	ACCATGAGAGGAGCCACGCGCCACTTTCAGGAAGCACTCTTATTCAGCAAGCGCTCC	4788
OY	4742	CAGAGCCGAAGCTGACGTTGACCGCTGTGAGATCGGCAAGTCCCGGCAAGCACCGTG	4801
Dp	4789	CAGAGCCGAAGCTGACATTCACCCCGGAGAGTCCGAAGTCCCAACACAGCATGTG	4848
OY	4802	CCCGAGACCAACCAACCCCATCTCGGCTTATGAGCACTGTCTTGGGGGCTGATGTGC	4861
Dp	4849	CCCGAGACCAACCTCACTCCATCTTCCCTTAATGAGCACTTGTCTCGGGGGCTGACTGT	4908
OY	4862	GTGACCTGTATCGAGCAACATCCCTCGCTTTCAGACCCCACTTCATACCCCGCGGC	4921

Db	4909	GTGGAACCTGTACCGTGTGCATCATCCATTGGACCTTTGACCCACCTCCATACCCCGAGG	4966
Oy	4922	ATCCCTCTGTA---CGACGCGCTGCTCACTACCTGACCCCGACACTTGGCCCCCAACCC	4978
Db	4969	ATCCCTCTGGAAGAGACGACGCGGACGCTACCTACCTGCCCCCGGACTTGGCCCCGAGCCC	5022
Oy	4979	ACCTAACCGGACCTGTATCCACCCCTACTCATTCGGCGGCTACCCCGACAGCGCGCGCTG	5033
Db	5029	ACCTAACCCACACTGTATCCCACTTACTCTATCCGCGGCTACCTGTACACGCGCGGCTGT	5088
Oy	5039	GAGAAACGGGACGACATCATTAATGAATACATCACTCCGACGAGATGATGACACCAACAG	5098
Db	5089	GAGAACCGGACGACATCATTAATGAATGAATCACTCCGACGAGATGATGACACCAACAGCT	5148
Oy	5099	GCCACCGGCACTGAGCCGACGACGAGCTGATATGTGTGAGGGGCTCTCGCCCCGAGAGTCTG	5158
Db	5149	GCCTCCGCAATGGCCGACGCGTGTGACATGTGTGAGGGGTCTGTACCGCGAGAGTCTTG	5208
Oy	5159	CTGGCACTCAACTACGCTGCGGGATCCCGAGGACATCATGACTCTGTCCAAATGTCACAC	5218
Db	5209	CTGGCCCTCAATATTATTCGCTGTGGCCCAAGAGGCAATTATCGACTGTGCCAATGTGCACAC	5268
Oy	5219	CTGGCTGTGTCTCGTGGCCCGGACACACAGGACCCGACGACCGCATGAGACCGGCTTGGC	5278
Db	5269	CTGGCCGTGTGTGTGACCAACACGCGACGACCCCTGTGACCGCATGACCGCTTGTGC	5328
Oy	5279	TACCTCCCAACCGGCGCCGACGCGCTTACGACGCGGACACAGCACTGCCACTCTCCCA	5338
Db	5329	TACCTCCCACTGTGGCCCCCACCCTTACGACGCGGACGACGATGATCACTGCTGCCCA	5388
Oy	5339	GAGAGTCCAAACAATTGACAAACCAACCAACGCTTCGTGTGAGCGGAGGAGAC	5398
Db	5389	GAGAGCCCACTACCTAGCTAAACCAACTGCCATCTTCATCGGAGCGGAAACG----	5448
Oy	5399	CGGATGAGAGCGGAGACCGGATTCGGGAGGGGAAAAATGCATCTTCACTCCACACAG	5458
Db	5449	-----GGAACGTGACCGGGAAGAAACAAGTCACTTCACGCTTACACT	5498
Oy	5459	ACGGTGAAGACGACCACTATGAGACTGTGTACAGACGAGACGCGGACGACGCGAC	5518
Db	5491	ACAGTGAAGATGACCACTTGTGAAGACTGTGTGAGACGAGACGAGCGGGGCTGGGGG	5558
Oy	5519	AGCAGCGGCGGGGTGGGGGAGGACGAGACCGGCCCGCTTCCACTCCATGCCACAG	5578
Db	5551	AGCAGCG-----CCCGCTCCCAACCAACAG	5588
Oy	5579	CACCTGCCCATCTCCCTCGGACCCAGATGCCCTCCAGACGACACCAAGTGTTCAC	5638
Db	5581	CACCTGCCCATCTCCCGGACCCAGACGCTTTCAGACGAGAGCCCAAGTGTGTGAC	5648
Oy	5639	AACACAGGCAATGAAGGGTATCATACGCGTGTGAGCCACAGACCCACGCTCTGAG	5698
Db	5641	AACACGAGCAATGAAGGGGTGTGTAACCTCGTGAAACCCGACGACCCAGGCTGTAG	5708
Oy	5699	TCCACTTCACTTCTCAACCCGTTTCGCGACGCTGCAATTCCACCTGCCACCACTGC	5758
Db	5701	TCCACTTCACTTCTTGGCTGTTCGCGACGCTGCAATTCCACCTGCCACCACTGC	5768
Oy	5759	CCACTGGGCGGACCTCTGATGGGCTTACCTTACCTCATGAGCCGCTTGTGCGCC	5818
Db	5761	CCACTGTGTGCACTTGAAGGGGTCTTACCTTACCTCATGAGGCCGCTCTGTATCC	5828
Oy	5819	AAGGAGGCCCCCGGGGTGCGCCGCGACGAGGGGCCCCGAGACACCGGCAATGCTTC	5878
Db	5821	AAGGAGACTTCTGGGTGGCCCGGCGGAGGGGCCCCGTGTGAAGCGTGGCCATATGCTTC	5888
Oy	5879	CTCGCAAGCCCCGAGCGCGCTCGGGCTGAGCGCGCTCTCTCCCGACGAAGGCTCG	5938
Db	5881	CTCAACCAACCCCCGGCGCG-----GAGGCGCGCTCTCTCAACCAAGAGCTCG	5938
Oy	5939	GAGCCCCGACCTAATGCTCTCTGTCTTGTGGCAAGCACTGAGCCCGACCCCTGCG	5998

Db 5932 GAGCCCGGATCCCTAGACACCCCGGAGCTCCAGCACAAGCATCGCCCGACCCCGACA 5991  
QY 5999 AAGAACTGCGACCTTACACCGGCAAGCCCGGAGCCCGGCGCACTGCTGGGCTCG 6058  
Db 5992 AAGAGCTTGGACCCCAACCAATCCGAGTCCGAGCCCGCGGG---CCCACTGGGCTCA 6048  
QY 6059 GACCCGACCGGGGAAAGACTCAAGATAAACCTTTTCCATCCAGGAATCGAACTCCGT 6118  
Db 6049 GATCTGCACCGAGAAAGACTCAAGATAAACCTTTTCCATCCAGGAATCGAACTCCGT 6108  
QY 6119 TCTCTGGGTTAACCA---GGCAGAGCTTACAGCCCGGAGGGGTGAGGCCCGTCAAGCCCT 6175  
Db 6109 TCTCTGGGTTAACCAAGTGAAGCTGCTACAGCCCGGAGTGGAGGCCCATCAAGCCCG 6168  
QY 6176 GTGAGCTCACCCAGCTTGAACCCAGCAGCAAGGGGCTCCCAAGCACTGGAAAGCTCGAC 6235  
Db 6169 GTGAGCTCCCCCGAGCTGACCCAGCAGCAAGGGGCTCTCCAAACCTTGGAAAGCTGAG 6228  
QY 6236 AAGAGCACCTGGAAGGGGAGCTGCGGCGCCAGCAGCAGGGCCCGTGAAGCTTGGCGGG 6295  
Db 6229 AAGAGCACCTTGAAGGGGAGCTGCGGCGCAGCAGCAGGGCCCATGAAGCTCAAGCGG 6288  
QY 6296 GAGGCGCGCACTTCCACACCTGCGCGCTGAGAGCCAGCCCTGTCAGCGCG 6355  
Db 6289 GAGGCTGCCCATCTCCACACTGCGGGCACTGCGCGAGGCAAGCCCTCATCAAGCCCA 6348  
QY 6356 CTGCTTCAAGCGCCCGGAGGGGTCAAAAGTCAACAGCGGGTGTCACTTGGGCCAGCAC 6415  
Db 6349 CTCTCTCAGACTGCGCCGAGGATCAAAAGTCAACAGGGGTGTCACTTGGGTCTAGCAC 6408  
QY 6416 ATCAGTGAAGGTATCAACAGAGCTACACCGGCAACACCCAGAGAGCTCAGCGACCC 6475  
Db 6409 ATCAGTGAAGGTATTAAGCAGAGCTACACCGGCAACACCCAGAGAGCTCAGTGGCCCC 6468  
QY 6476 CTGCGCGCGCCCTCTACTCTTCTCTGCGGCGCAGCTGCGCCCTCTGGAACCTCGCGCG 6535  
Db 6469 CTTCGCGCGCCCTCTACTCTTCTCTGCGGAGCAGCTGCGCTGTGGATCTTCGCGCG 6528  
QY 6536 CCAACCAAGTACTTCACTCTCCGCGCGCGAGCAATGATGTCGCGCGCTGCTCCCG 6595  
Db 6529 CCAACCAAGTACTTCACTCTCCGCGCGCGAGCAATGATGTCGCGCGCTGCTCCCG 6588  
QY 6596 CACAGGAAAGGGGCGAAGAGGTCTCAGAGCGAAACAGAGCTGGGTGGGTGGT 6655  
Db 6589 CACAGTGAAGGGGCGAAGAGGTCTCCAGAACCCAGCAAAACATGGGCTTGGGCGAGAT 6648  
QY 6656 GAGGACGGAATTGGAACCTGTGTCCCAACCGAGGGGCTGAGCGAGCGACATCCCGG 6715  
Db 6649 GAGGATGGCATTTGAGCTGTGTCCCAACCGAGGGGCTGAGCGAGCAATGCTCGG 6708  
QY 6716 AGTGTGTGTATCCGCTGTGTATCCGGAGTGGGAAACAGCGAGCCCAAGAGATGGG 6775  
Db 6709 AGGCTGTGTATCCCACTGTGTATCGAGAGGGGAAACAGGGCGAGCC---CAGAGTGGG 6765  
QY 6776 TCCAAATCTCCAGGCAACCAAGCGCGCGCAGCCCTTCTTCAAGCAAGCTGACCGAGG 6835  
Db 6766 TCTAAATCTCCAGGCAACCAAGCGCGCGCAGCCCTTCTTCAAGTGAAGCTGAGAGG 6825  
QY 6836 AACTCCGCAATGTCAAGTCCAGAGAGCAGAGATCAACAAGAGCTGAACCCCAAC 6895  
Db 6826 AACTCCGCAATGTGAAGTGAAGAGCAGAGATCAACAAGAGCTGAACCCCAAC 6885  
QY 6896 CCGAATGAGCTGAATTAATATCAATATCAAGCAGCTGGGAGCGAGATCTTCAATATGCCG 6955  
Db 6886 CCGAATGAGCTGAATTAATATCAATATCAAGCAGCTGGGAGCGAGATCTTCAATATGCCG 6945  
QY 6956 ATCAACGGAACAGGCGCTTATGAAGCTTATGAAGCAGGCGGTGAGAGAACTGCAAGACC 7015  
Db 6946 ATCACTGAAGCAGGCGCTTATGAAGCTTATGAAGCAGGCGGTGAGAGAACTGCAAGACC 7005  
QY 7016 AACATGGGCTGAGGCGCATATTAAGAAAGCACTCATGGGTAAATATGATCAGTGGAA 7075  
Db 7006 AACATGGGCTTAAAGGCAATTAATTAAGAAAGCACTCATGGGTAAATATGATCAGTGGAA 7065

QY 7076 GAGTCCCGCGCTGAGCGCAATGCTTTTAAACCTCTGAATGSCAGTGCAGCCTGCC 7135  
Db 7066 GAGCCCCCGGCTGAGCGCAATGCTTTTAAACCTCTGAATGSCAGCCTGCC 7125  
QY 7136 ---GCTGTATGGCCATTAACCGCTGTGACGAGCGGATGACCACTCATCTGGCA 7192  
Db 7126 GCTGTGTATGGCCATTAACCTGTGACGAGCGGATGACCACTCATCTGGCA 7185  
QY 7193 GGTGCGCGGGAAGGCAAGTCTTGGGCAACCCAGCAGCCGAAAGCAAGTCCCG 7252  
Db 7186 GGTGAGGTGGGAAGCAGAGTCTTGGGCAACCTTGAAGCAGCCGAAAGCAAGTCCCG 7245  
QY 7253 GCGCGGGGCTGGCATGTGGGAGCGGCGCACCTGTGTCTCTCAGTGCATCGAGAG 7312  
Db 7246 GCAACAGGCTTACGTCGAGAGCGAACCCCTTCTGTCTCTCAGTGCATCGAGAG 7305  
QY 7313 GACTGCAACCGCGGAGCGCGCTCAACCAACCGGTGTGGAGAGCAGGCTTGTCCGA 7372  
Db 7306 GACTGCAATCGCGGAACACACTCAACCGGTGTGGAGAGCAGGCTTGTCACTGCA 7365  
QY 7373 GGTTCAGCGCATTTCCCTTAACAACCTCTGATCATGCGCTGAGCGGCTCATGGCT 7432  
Db 7366 GGTTCAGCGCATTTCCCTTAACAACCTTGTATGAGGCTACAGGCAAGTGTCAAGCC 7425  
QY 7433 TCCCGACCCCGACCGGCGCTCCCGGAGGAGCGGCGCTGAGCGCCCGACAGCGC 7492  
Db 7426 TCCCGCGCCCGACCTGAGCTTGGGAGGAGGAGGCGCTTACTGTGCCACAGCGC 7485  
QY 7493 TGGGACGAGAGCGCCAGGCACTGTCTGTGCGAGTACAGACACTTCCGACAGAG 7552  
Db 7486 TGGGATGAGAGGCGCCAGGCACTGTGTGTTCACATTAAGAGCACTTCGAGCAGGAG 7545  
QY 7553 TGACTCAAAACAGGCGGGGAGGCGGCGGTGEBERSFRLUTGAGGTCCAGGAGC 7612  
Db 7546 TGAACACGATTTGGGGGAGAGCGG-----TGCAGGTCCCGACAA 7586  
QY 7613 CACAGGAAGGCGCTGACAGAGGCGGCGTGCAGACTCCCGCAACCAAGAAAGAGCC 7672  
Db 7587 GGCAGAAAGCGCCACAGCAGAGAGCAGACAGCTGTACTTCGAGACTGAGAAAGAGCC 7646  
QY 7673 CCTGAGTCCGCGCTGC-----GCTTCATCATCTGTCTGTCAGAGCGG 7717  
Db 7647 CCGAGTCTGCGCTGCGCTGCATCCGTCGTCGTCATCTGTCATCTCAGAGCTGG 7706  
QY 7718 CATCTTGCCTGTCTTAAAGCTTAAAGACTCCCGCGGCTGCGCTGTGAG-- 7775  
Db 7707 CATCC-TGCCTGTCTAAAGCTTAAAGACTCCCAACCGGCGTGGCGCTGCGAGTG 7765  
QY 7776 --ACCTTACTCAGGGAGTGTAACTGTGTGCTGCGGAGAGGAGGGAA----- 7821  
Db 7766 ACCTTAACCTCAGGGAGTGTAACTGTGTGCTGAGAGGGAGTGAACAGAGGGAGGG 7825  
QY 7822 -----GGGCGGGGAGGGGAGCGGCAAGCGGTGTGGCAGCCACACAGCGCGGCAAGG 7876  
Db 7826 ACAGCGGGCGCAAGAGGGGGGAGCAGCAATCGTGTCTAGTGGCACTCGTGAATCGGG 7885  
QY 7877 CCGCAGGAGCCCAAGAGATGACAGCACTTCAAGCGCACTGCT--CTCCGAAT 7934  
Db 7886 TCAGCGGGAGCCACACAGAGCTGACAGGCACTCAATGCGCACCGCTCGCGCTTAC 7945  
QY 7935 GCAATTTGGAACCAAACTCTAACTGAGCTGCAAGCCCGCGCGCTCCCTCGCGCTCCCA 7994  
Db 7946 GCAATTTGGAACCAAACTCTAACTGAGCTGCAAGCTCGCGGTCTGCTCTGCGCCCA 8005  
QY 7995 TCCCGCTTAGCGCTTGAACAGTGAAGCAGGCGCTGTTCAGCGCCCGCAGTGGCTGT 8054  
Db 8006 GCGCGCTTGTGCTTGAACAGAGAGTTCAGAGCTTATCTGCTTAAATGCTGTATCA 8065  
QY 8055 CCGGTCCCAACAGACTGCGCCAGCCCAAGAGATGCTGGAACCAAGTCAAGGCGAGTGG 8114  
Db 8066 TCGCACTTCAAA---GGCAGCCAGCCCAAGAGCTGGAGGCCCATGAGCAAGTGG 8122

QY	8115	GCGGCAAAAAGGGCAGGTGGGGCTGGGGGAAACGANTGCCTCGAAGGACTGGACTGTT	8174
Dd	8123	GTGACACAAGGGGGCTGG-----CTGGGGCACGANtGCTTGCAGGAACTGACCGTT	8174#
QY	8175	TTTTCAACATCGTTGCCGACAGCGGTGGGAAGAAGGACAGATGTMAATTGATGTGTGGT	8234#
Dd	8175	TCCGGCGCTGTGGTCGTGGCAACGG---GAGGGAAAGCAcGTGTMAATG--GTGTTGGC	8228#
QY	8235	TTTAACGGTATATTTTGTGATACTTCATGMAATTAATTCAAGATGTTTAAcGACGAAGG	829#
Dd	8229	TTTAACGGGTATATTTTGTGATACCTTCATGAGTTAATTCAACCGCTCACA.CAAGAAGG	8288#
QY	8295	ACTTACCAGATTACTACTGCTCTGNGCTTTGATCTGTGCTAACGGTCAAGAAGCGNCT	8354#
Dd	8289	ACTGGCCC-GTGTTTCTCCCGCTGGC--TTTGGTCTTAACCTA.CTGTTTGAGAAGCACT	8346#
QY	8355	GCAGGCGGACAGTCGGTAGACCCCATCACTGSCAGGACCAAGGGGGCGGGAC.TGCTCGTC	8414#
Dd	8347	GC---CAGCC.AAGGCGGTGGCCCA.CAATACGACAGACTTGGGGGGG.CAAGGGCGCTCGG.CAC	8403#
QY	8415	ACGGCCCGCTGTGCTCCCTCCCTCCCTGCTGGGAGAAATGAATCGANtGCGTATTC	8474#
Dd	8404	ACGGCACTGTG---CCCTTCCCA.CACCTTA.CTTAGGAAATGAC.TTGAAGCGTATTC	8460#
QY	8475	TGTGGCGGCATTTTTCGACAGGGGTGGTGTATTTCTCATTTTACA.CCGTCGTTCTAATTA	8534#
Dd	8461	TGTGGCGGCTCTCTGTGCA---CGCGGCATTTCTGTCA.TTTTACACATGTTGTTCOAATTA	8517#
QY	8535	AAAAAGCGAAT	8544
Dd	8518	AAAAAGCAAAT	8527

```

RESULT 12
US-09-819-104A-6
? Sequence 6, Application US/09819104A
? Publication No. US20030027137A1
? GENERAL INFORMATION:
? APPLICANT: Chen, J. Don
? TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
? TITLE OF INVENTION: AND USES THEREOF
? FILE REFERENCE: UMG-030
? CURRENT APPLICATION NUMBER: US/09/819,104A
? CURRENT FILING DATE: 2001-03-27
? PRIOR APPLICATION NUMBER: 60/1193,138
? PRIOR FILING DATE: 2000-03-29
? NUMBER OF SEQ ID NOS: 6
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO: 6
? LENGTH: 7386
? TYPE: DNA
? ORGANISM: Mus musculus
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1) ..(7386)
? US-09-819-104A-6

```

Query Match	56.0%	Score 4797;	DB 3;	Length 7386;
Best Local Similarity	80.3%	Pred. Nis. 0;		
Matches 6118; Conservative	0;	MissMatches 1200;	Indels 301;	Gaps 30;

QY	2	ATGCGGGCTTCACACAGCTGTGGACACAGCGTGAGAGGCGCATGAGCCCGGTACCG	61
Db	1	ATGTCGAGATCCACACAGCTGTGGACACAGCATGGCGGGCTGTGAGCCCGGTACCCA	60
QY	62	CCCCACAGCCCTTTCCTTCCGACGTGCATATGCGCCCGAGACGACAGGAGCTGGGCTCTG	121
Db	61	CCCCATGACATCTCTTACCCGGTGAAGTAAACCCGGTCCACAGGAGTGGGGCTGT	120
QY	122	GAGTACCAAGACCACTCCCGGACTATGCTCCACCTGTGCGCGGGCTCCATCATCCAG	181
Db	121	GAGTACCAACACACCCCGGTGATTAACCTTCACACCTGTACACCGGTTCCATCATCCAG	180

QY	182	CCCCAGCGCGGAGAGGCCCTCCCTCGTGTCTAGATTCCAGCCCGGGAAATGAACGGTCCAG	241
Db	181	CCACAGAGAGAGCGGCCCTCACTGTGTTCAGAGTTCCAGCCTGGAGTGAACGGTCTTAG	240
QY	242	GAGCTCCACCTGCGGCGCAGAGTCCCACTCAATACCTGCCGAGCTGGGAAATCAGAGATG	301
Db	241	GAGCTCCACCTGCGGCCCTGAGTCTCCGCACGTTCTGTGCTGAGCTGGGAAACCCGACATTA	300
QY	302	GAGTTCAATTGAAAGCAGCGCCCTCGGCTAGAGCTGTGCTGACCCCTGTGCGACCG	361
Db	301	GAATTCAACCGAGAGCAGACGCCGCCCGCTGGAGGCTAACCCGATACCTCTGTGCGCCA	360
QY	362	TCACCCCTGCTGGGCCACCGGCGCAGCTCGCGGATCTGAAAGCTTCAACAAAGACCTTAGC	421
Db	361	TCACCCCTGCTGGGCCACTGGGCACTGGGCAAGTGGGTCTGAAAGCTTTAACAAAGACCTTAGC	420
QY	422	CTGACGGGCAAGCTGGAACCGGTGTCTCCCCACAGCCCCCGSCACATGACCTTAGCTG	481
Db	421	CTGGCAGGCAAGCTGGAGCTGTGTCACTCTCCAGTCCCCCGCACGCTGACCTTAGAGCTTA	480
QY	482	GAGCTGTGCGGCCACCGCTGTCCAAGAGGAGCTGATCCAGAAC--TGGACCGCGTG	538
Db	481	GAGCTGTGCGGCATCTGCAGCTGTCCAAAGGAGGCTGATCCAGAACAGATTGGAACGCGCGTG	540
QY	539	GACCGAGAGATCAACATGTGTAGACAGAGATCTCTAAGCTGGAAGAAACACAGACAAAG	598
Db	541	GACCGTAGATCAACATGTGTAGACAGAGATCTTCAAGCTGAAGAAACACAGACAAAG	600
QY	599	CTGGAGGAGAGGCTGCCAAGCCGCGCGAGCTTGAGAAAGCCGTGTCAACCGCCGCATC	658
Db	601	TTGGAGAGAGAGGCCGCGCAAGCCGCGCGAACCCGAGAGCCGTGTGTGCGCACACCCATA	660
QY	659	GAGTCGAAGCACCGCAGGCTGTGTCAAGATCATCTACGACGAAACCGAGAGAGGCTGAA	718
Db	661	GAATCAAAAGCACCGAAGCTGTGTCAAGATCATCTACGATGAGAAACCGGAAAGAAACCGAA	720
QY	719	GCTGCACTCGGATTTCTGGAAGGCTTGGGGCCCCAGAGTGGAGCTGCGCTGTACAACAG	778
Db	721	GCCGCAACACCGGATCTTAGAAGGCTTGGGGCCCCAGAGTGGAGCTGCTCTGTACAACAG	780
QY	779	CCCTCCGACACCGGCAAGTATCATAGAAACATCAAAATTAACCAAGGCGCATGCGGAAG	838
Db	781	CCGTCTGACACACGCAAGTACCATGAATAACATCAAAATTAACCAAGGCGCATGCGGAAG	840
QY	839	CTAATCTTTGACTTCAAGAGAGGAAATCACGCTCGGAAACATAGAAAGCAAGATTCTGCG	898
Db	841	CTGATCTTTGTACTTTAAGCGSAGGAACACCGCGGCAAGCATGGGAACACGCTTCTGCG	900
QY	899	CAGCGCTATGACCAAGCTCATAGAGGCTTGGAAAAAAAGTGGAGCGCATCGAAAAAC	958
Db	901	CAGCGCTATGACCAAGCTCATAGAGGCGTGGGAGAAAGTATGAGCGCATAGAAACAAT	960
QY	959	CCGCGCCGCGCGGCGCAAGAGAGCAAGTGGCGAGTACTACGAAAAAGCATTTCCCTGAG	1018
Db	961	CCGCGAAGAGAGAGGCGCAAGAGAGCAAGTGGGAGTACTACGAAAAAGCATTTCCCGAG	1020
QY	1019	ATCCGCAAGCAGCGCGAGCTCAAGAGCGCATGACAGACAGGAGTGGCGCAGCGGGCAGT	1078
Db	1021	ATCCGCAAGCAGCGCGGAGCTCAAGAGCGCATGACAGACAGGAGTGGCGCAGCTGGCAGT	1080
QY	1079	GGGCTGTTCATGTGCGCGCGCCGACGAGACAGAGGTTCTCAGAGATCATGATGGCTCT	1138
Db	1081	GGGCTGTTCATGTGCGCGCGCCGACAGTGAAGCATAGAGTTTCTGAGATCATGATGGCTTG	1140
QY	1139	TCAGAGCAGGAAACTTGGAGAGACAGATGGCGCAGCTGTGCGGTGATCTCCGCCATGCTG	1198
Db	1141	TCTAGAGCAGGAAACTTGGAGAAACAGATGGCGCAGCTGTGCGGTGATCTC--GCCATGTTG	1198
QY	1199	TACGACGCTGACGAGCGCATGAGTCACTCAACATGAACCGGCGTTATGCGCGACCC	1256
Db	1199	TACGACCG--GACCAAGCAGAGGATCAAGTTCACTCAACATGAATGATCTCATGATGACCCC	1257
QY	1259	ATGAAGGTGTACAAGACCGCCAGGTGATGAACTGTGAGTGAACGAGAAAGAGACC	1318

1258 ATGAAGTCTACAAAGACCGTCAAGTTACCAACATGTGAGCAGCAGAGGACAC 1317  
1319 TTCGGGAAAGACTTCATGCAAGCATCCCAAGAACTTTGGCTGATGCCATCATCTCGAG 1378  
1318 TTCCTGAAAGATTATGACGACCTTAAGAACTTTGGCTGATGCCATCATCTCGAG 1377  
1379 AGGAAGACAGTGGCTAGTGGCTCTCTATTACTACCTGAAGAAATGAGAACTAT 1438  
1378 AGAAAGACGGTGGCTAGTGGCTCTCTATTACTACCTGAAGAAATGAGAAATTAAC 1437  
1439 AAGAGCTGTGTGAGCGAGACTATCGCGCGCGGCAAGAGCCAGCAACAACAGCAG 1498  
1438 AAGAGCTGTGTGAGCGAGACTATCGCGCGCGGCAAGAGCCAGCAACAACAGCAG 1497  
1499 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1558  
1498 CAACAAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1545  
1559 AAGAGATGAG 1618  
1546 AAG 1605  
1619 AAGCAG 1678  
1606 AAG 1665  
1679 GAG 1738  
1666 GAG 1725  
1739 GAGCAGATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1798  
1726 GAGCAGATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1785  
1799 CAGAGCGCGAGCTGCTCTCATAGAGCTGATGAGAGTTCTCGTGAAGAGAGAGAG 1858  
1786 CAAAGTTCAAGAGCTGCTCTCATAGAGAGTTCTCGTGAAGAGAGAGAGAGAG 1845  
1859 ATGGAAG 1918  
1846 ATGGAAG 1905  
1919 ATGAGTGGCTCCAAAGCTGTGTGAGAGTGAAGAACTTCACTCACTCAAGAGAG 1978  
1906 ATGAGTGGCTCCAAAGCTGTGTGAGAGTGAAGAACTTCACTCACTCAAGAGAG 1965  
1979 CAGAACTCGATGAGATCTTTCAGAGCAGCAGAACTGAAGTGAAGAGAGAGAGAG 2038  
1966 CAGAACTCGATGAGATCTTTCAGAGCAGCAGAACTGAAGTGAAGAGAGAGAGAG 2025  
2039 CAG 2098  
2026 CAG 2085  
2099 GTGAG 2158  
2086 GTGAG 2145  
2159 GAGGCTGAGAGCTTCAATGCTCTGAGAAATGAGAGTCCAGAGAGAGAGAGAGAG 2215  
2146 GAGGCTGAGAGCTTCAATGCTCTGAGAAATGAGAGTCCAGAGAGAGAGAGAGAG 2205  
2216 CAGAGCAGTGTCAACAGAGCTCAGACCGAGAGAGATCCCTCTCTCACTGAGAGCC 2275  
2206 CAGAGCAGTGTCAACAGAGCTCAGACCGAGAGAGATCCCTCTCTCACTGAGAGCC 2265  
2276 GCGAG 2335  
2266 AGGAG 2316  
2336 CCGCAG 2395

2317 ACCCAGCAGCTGTCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2376  
2396 GCGTCAAG 2455  
2377 GTTCTGTATGAG 2433  
2456 CCTCTGTGTGTCTCCAG 2515  
2434 CCGGCTACTGTGAG 2493  
2516 GAGGAG 2557  
2494 GATGCAAG 2553  
2558 GTGAGCAG 2617  
2554 GAGGCTCTGAG 2613  
2618 GAGGCTCTGAG 2674  
2614 GAGGCTGAG 2673  
2675 AAGGAG 2734  
2674 AAGGAG 2730  
2735 CCGCAG 2794  
2731 ACCCAG 2790  
2795 GCGCAG 2854  
2791 GGTGAG 2850  
2855 CCGGAG 2914  
2851 CCGGAG 2910  
2915 GCGATGAG 2974  
2911 GCGATGAG 2967  
2975 ACCCAG 3034  
2968 CCAAG 3027  
3035 CCTCAG 3094  
3028 TCTCAG 3087  
3095 AAG 3131  
3088 AAG 3147  
3132 GAG 3191  
3148 CTGAG 3202  
3192 AGGAG 3251  
3203 AGGAG 3256  
3252 TGCCTGTGAG 3311  
3257 TGCCTGTGAG 3313  
3312 ACCGAG 3371  
3314 ACCGAG 3373  
3372 CCATCTCCAG 3431  
3374 CCATCTCCAG 3430

QY 3432 CGGTGGGCGCTGTCAACATGGGGCTGGCCCTGAGATGGAACCCCAAAAGCTGGACCTT 3491  
| | | | |  
Db 3431 CCATGGGGCGCTCTCAACATGAGAGCTGCCCTTGCGGTGACCTTAAGAGCTG----- 3483  
| | | | |  
QY 3432 TCAGCGAGATGAAGAGAGAGAGAGCTGTCCCAACGGGGCCAGGCTGGGCCACCGAGAGGCC 3551  
| | | | |  
Db 3484 ----- 3483  
| | | | |  
QY 3552 TGGGGGTGCCCAAGCCAGAGAGCGTCCGTGTGAGAGGAGCAGCTGTGGCTCAGTTTC 3611  
| | | | |  
Db 3484 -----GGAGACGACT--GGCTCCGCCA 3504  
| | | | |  
QY 3612 CGGGCGAAAGCATCAACAAAGGCATTCCAGACAAGGGTGCCTTGAGCAGGCCCATCA 3671  
| | | | |  
Db 3505 CCAGTGGAAGCATCAACAAAGGGCTCCCGATACCGGGGCTGCAGACGG-----CCCCA 3557  
| | | | |  
QY 3672 CATACCCGCGCTCCATCAACCCAGCGACGCGCAGCTGACGTCTTGTAACAAGGGCACCATCA 3731  
| | | | |  
Db 3558 GCTACAGAGGCTCTATCAACCCAC-GCACGCCGCGAGACGTCTCTTAACAAGGGTACCATCA 3616  
| | | | |  
QY 3732 CCAGATCATCGGCGAGAGCAGCCCGAGCTTGGACCGCGCGCGGAGAGCAGCGTTCG 3791  
| | | | |  
Db 3617 GCAGATCGTCCGTGAGAGACAGCCAAAGTCGCTTGACCGGGCAGAGAGACACCTTC 3676  
| | | | |  
QY 3792 CCAAGGGCCAGTCATCTACGAAGCGAAGAGGCGCCAGCTTGTGTCTATGAGGGTGGCA 3851  
| | | | |  
Db 3677 CCAAGGGCCAGTCATCTATGAGGGCAGAAAGGCCAGCTCTATCTATGAAAGGTGGTA 3736  
| | | | |  
QY 3852 TGTCTGTACCCAGGTCTCCAGAGAGAGCGAGAAAGCATCAGAACCCCCCATGAGA 3911  
| | | | |  
Db 3737 TGTCTGTGTCAAGTGTCTTAAGAGAGATGAAAGGAGAGCTCGGCGCCACCCCATGAGA 3796  
| | | | |  
QY 3912 CGGCGCGCCCGCAAGCGCAGCTATGACATGATGAAGGGCGCGTGGGACAGGCAATCTCT 3971  
| | | | |  
Db 3797 CTGCGCGCCCTTAAGAGCAGCTATGACATGATGAGAGGGCGGTGTAGGACAGACTTCACCT 3856  
| | | | |  
QY 3972 CAGCCAGCATGAGAGTCTCATGGGCGGTGCATTCGCCCGAGCGACAGCCCCCACC 4031  
| | | | |  
Db 3857 CAGCCAGCATGAGAGGAGCTCATGGGCGGCGCATCCCTGAGCAGC-----ACAGCCCC 3910  
| | | | |  
QY 4032 ACCTCAAGAGAGAGACCAACATCCGCGGTCATCAACAAAGGATCCCTCGGTCTTAG 4091  
| | | | |  
Db 3911 ACCTCAAGAGAGAGATCAATCCAGGCTCATCAACGAAGGCATCCCGAGGTCTTAG 3970  
| | | | |  
QY 4092 TGAAGGCAAGAGAGACTACCTGCTCGGAGAGCCAAAGCTCTTAAGCGGAGGCGACGC 4151  
| | | | |  
Db 3971 TGAAGGCGAGAGAGACTTACCGCGGAGAGCCAAAGCTCTTGAAGCGAAGAGGACAC 4030  
| | | | |  
QY 4152 CTCGCGCCCGACCGCTTCACGAGGACCTGACCGAGGCTTAAGAAGCGCAGGCC----- 4204  
| | | | |  
Db 4031 CACTCTCCCGACCAACCACTCGGAGACCTGACCTGAGACTTAAGGCCCGGCGCTGAGAC 4090  
| | | | |  
QY 4205 --CTGGGGCCCCCTGAAGCTGAAGCCGCGCCATGAGAGGCTGTGTGCGCACAGGTGAAGAG 4262  
| | | | |  
Db 4091 CTCTGGGTCTCTGAAGCTGAAGCCGACCTCAAGAGGTGTGTAGCAATCTGGAAGAGAG 4150  
| | | | |  
QY 4263 CGGGCGGCTCATCATGATGATCCGCGGAGAGAGTGGCGGACAAGCCCGAGCTGCCCG 4322  
| | | | |  
Db 4151 CGGGCGGCTCTATCATGAGATCCCGAGAGAGAGTGGCGCGCACACTTGAGCTTACCC 4210  
| | | | |  
QY 4333 TGGCCCGCGCGCTCAAGAGAGGCTTCATCAACGAGGCGACCCCGCTCAAGTACGACA 4382  
| | | | |  
Db 4211 TGGCACACAGGCGCTCTGAAGGAGGGTTCATCACCAAGGCGACCCCACTCAAGTACGACT 4270  
| | | | |  
QY 4383 CCGGGCGCTTCACACTGGCTCCAAAAGACAGAGTACGCTTCCTCATGGCAGCCCG 4442  
| | | | |  
Db 4271 CTGGGCGACCTTCACCTGGCACAAGAAACAGAGGTGGCTCATCATTCGCGACCCCG 4330  
| | | | |  
QY 4443 GCCGAGCGTTCACCGCGTGCACCGCGTGTGATGTGGCGAGCGCGGCGGACCTGGAAC 4502  
| | | | |  
Db 4331 GCCGCGCTTTCCTGCTGACCGCGCTGAGCATTAATGCTGACGCGCGGAGACTGAGAC 4390  
| | | | |

QY 4503 GTGCTGCTACAGAGAGAGCTGAAGAGCCGGCCAGGAGCGGCGAGAGCTCGGGGGCT 4562  
| | | | |  
Db 4391 GTGCTGCTATGAAGAGAGTGTGAAGAGCCGGTGAAGGAGCCAGGAGTGTGAGGGGGCT 4450  
| | | | |  
QY 4563 CCATTGGCGGGGGCGCCCGGTCATTGTCGTGAGCTGGTAAAGCGCGGAGAGAGCCCG 4622  
| | | | |  
Db 4451 CCATCAACAGTGGGGCTCAAGTGTGTGCTGAAGTGGCAAGCAGGCAAGCCACAC 4510  
| | | | |  
QY 4623 TGAACCTATGAGAGCAGCGGGGACCTTTTGCGGCGACCTCCACAGAGTTGCGCGTGA 4682  
| | | | |  
Db 4511 TGACTTTAAGAGAGCAGCGGGGACCTTACACAGTCACTGCAAGTGGCTCCCTGGA 4570  
| | | | |  
QY 4683 CCATCGGGAGCCACCGCGCGCTGACAGAGGGCAGCTTTGTCAGCAAGGATCCC 4742  
| | | | |  
Db 4571 CCAAGAGGAGCGCCAGCGCCTTCAGGAAGGAGCGCTCTATTCAGCAAGGCGTCCC 4630  
| | | | |  
QY 4743 AGGACCGAAAGCTGACGTGTGACGCTCGTGTGATGGCAAGTCCCGGACAGACCGGTGC 4802  
| | | | |  
Db 4631 AGGACCGAAGCTGACATCTACACCCGGAGATGCGAAGTCCCAACAGACTGTGC 4690  
| | | | |  
QY 4803 CCGAGCACCAACCAACCCCATCTGCGCTATGAGCAGCTGTGGGGCGTGAATGGCG 4862  
| | | | |  
Db 4691 CCGAGCACCAACCTCAACCCCATCTGCCCTATGAGCAGTTGCTCCGGGCGTGAATGGG 4750  
| | | | |  
QY 4863 TGAACCTGTATGCGAGCCACATCCCGCTTGCTGACCCCACTTCATACCCCGGCGCA 4922  
| | | | |  
Db 4751 TGAACCTGTATCGTGTGATCCCATTTGGCTTTGAGCCCACTTCATACCCCGAGGGA 4810  
| | | | |  
QY 4923 TCCCTGTGGA---GCGAGCGGTGCTACTACCTGCGCGGACACCTGCGCCCAACCCA 4979  
| | | | |  
Db 4811 TCCTCTGAAAGCACAGCCGAGCTTACTACCTGCGCGGACCTTGGCCCCCAGCCCA 4870  
| | | | |  
QY 4980 CCTACCGCACCTGTACCCACCTTACTCATTCGCGGCTACCCCGACAGCGGCGGCTTGG 5039  
| | | | |  
Db 4871 CCTACCGACACTGTATCCCACTTACTCATTCGCGGCTTACCTGACAGCGGCGCTTGG 4930  
| | | | |  
QY 5040 AGAAGCGGAGACCATCATCATGATGATCACTCGGAGCATGACCAACCAACAGG 5099  
| | | | |  
Db 4931 AGAAGCGGAGACCATCATCATGATGATCACTCGGAGCATGACCAACCAACAGGCTGG 4990  
| | | | |  
QY 5100 CCACCGCATGGCGGAGAGCTGATATGCTGAGGGGGCTGTGGCGCGGAGTCTGGC 5159  
| | | | |  
Db 4991 CTTCCGCATGGCCAGGCTGTGACATGTGAGGGGCTGTGACCGGAGAGTCTGGC 5050  
| | | | |  
QY 5160 TGGCACTCAACTACCTGCTCGGGTCCCGAGGATCATGACCTGTCCCAAGTGCACACC 5219  
| | | | |  
Db 5051 TGGCCCTCAATTAATTCGCTGGCCCAAGAGGATTAATGACCTGTCCCAAGTGCACACC 5110  
| | | | |  
QY 5220 TGCTGTGCTGTGTCGCCGACACCAAGGACCCCAAGCAGCGCATGAGCGGCTTGGCT 5279  
| | | | |  
Db 5111 TGCCGTGTGCTGTGTCACCAACGAGGACCCCTGCGACCGCATGACCGGCTTGGCT 5170  
| | | | |  
QY 5280 ACCTCCCGACCGCGCGCCAGGCTTACAGACGCGCAAGAGCTCCCACTCTCCAG 5339  
| | | | |  
Db 5171 ACTTCCCACTGCGCGCCCACTTACAGACGCGCAAGTACGCTCAAGCTGTCCAG 5230  
| | | | |  
QY 5340 GAGGTCCAACACTTGAACAAACCAACCAAGTCTCTGTCGAGGGGAGCGAGACG 5399  
| | | | |  
Db 5231 GAGGCCCACTCACTAGCTTAACCACTGCAATCTTCACTGAGAGGG----- 5280  
| | | | |  
QY 5400 GGGATCGAGAGCGGAGCCGGAGTGGGAGCGGAGAAAGTCAATCTCTCAAGTCAACGAG 5459  
| | | | |  
Db 5281 -----GAACGGGAAAGTGAAGCGGGAAGAGCAAGATCATCTCTCAAGTCACTA 5332  
| | | | |  
QY 5460 CGGTGAGACAGCAACCACTGTGAGACTGTGTAAGAGAGAGAGCGGCGAGCGGCA 5519  
| | | | |  
Db 5333 CAGTGAAGCATGACCACTGTGAGACTGTGTAAGAGAGAGAGCGGCGTGGGGGCA 5392  
| | | | |  
QY 5520 GCAAGCGCGGGGAGTGGGGGAGAGCAGACCGCCCGCTTCCACTCCGATGCCACAGC 5579  
| | | | |  
Db 5393 GCAAGCG-----CCCGGCTTCCACACCAAGCAGC 5422  
| | | | |  
QY 5580 ACTGCGCATCTCCCTTCGAGACCAAGATGCCCTTCAGAGAGACCAAGTGTGTTCA 5639  
| | | | |



Db	5423	ACTGCGCCATCTCCCCCGGAGCCAGAGACGCTTGGACAGAGGCGCATGTGTCTGACAA	5482
Qy	5540	ACACAGCGATGAGGGATATCATCAACCGCTGTGAGGCCAGAGCCCACGGTCTTGAGGT	5699
Db	5483	ACACGAGCATGAGGGGCGTGGTCAACCTCCGTGGAACCCGGCAGCGCCACGATCTCGAGGT	55422
Qy	5700	CCACTCTCACTCTCTCAACCGGTGTGGCCGACGTGCACATTCCCACTGACCACTAATGCC	5759
Db	5543	CCACTCTCACTCTTTCGTGCTGTCCGCCACAGCTGGCCACATTCCTCACTGACCACTAATGCC	56022
Qy	5760	CACGTGGGCGGACCGCTGATGAGGGGTCTACCCCTACATGAGAGCCGCTTTCGTCGCCA	5819
Db	5603	CACCTTGGTGGCACTCTTGAAGGGGTCTACCTTACCTCATGAGCCGCTCTGTTCACCA	5662
Qy	5820	AGGAGGCGCCCCCGGGTTCGCCCGGCAGAGCGGCGCCCGACAGACACACCGGCATGCTTCC	5879
Db	5663	AGGAGACCTCTTCGGGTTCGCCCGGCAGAGCGGCGCCCGTGTGAGAGGTGGCATGCTTCC	57222
Qy	5880	TGCGCCAGACCCCCCAGCGCCGCTTCGCGGAGTGGAGCCCGGCTCTCTCCCGCAGCAAGGCTTCG	5939
Db	5723	TCACCAAAACCCCGGCGCCG-----GGAGCCCGCTCTCTCAACCCAGCAAGAGCTTCG	5773
Qy	5940	AGCCCCCGGCGCCCTTATGTGCTCTCTCTTCGCGCACGCCACATGCGCCCGCACCCCTGCGA	5999
Db	5774	AGCCCCGATCCCTTATGCAACCCCCAGCTCAACATCAACAGCCATGCGCCCGCACCCAGCA	5833
Qy	6000	AGAACCTTCGCACTTCACCAACCGCACCGCCGCGCGCCGACCTTGCCTCGGCTTCG	6059
Db	5834	AGAGCTTTGCAACCCCAACATGCTGATCCGGACCCCGCGGG-----CCACCTTCGGCTCAG	5890
Qy	6060	ACC CGCACCCGGAAAAAGATCAAAAGTAAACCTTTTCATCAGAGAACTGGAATCCGCTT	6119
Db	5891	ATCTGCAACCGAGAAAAAGATCAAAAGTAAACCTTTTCATCAGAGAAATGGAACTCGCTT	5950
Qy	6120	CTCTGGGTATCAC---GGCAGCAGCTTACACCCCGGAAGGGGTGAGAGCCCTCAGCCCTG	6176
Db	5951	CTCTGGGTATCACAGTGAAGCTGAGCTGAGCAGCCCGGAAGGGGTGAGAGCCCTCAGCCCG	6010
Qy	6177	TGACTCTCAACCAATCTGACCCACGACAAGGGGGTCTCCCAACCACTGGAAAGCTCAGCA	6236
Db	6011	TGAGTCTCCCCCAGCTGACCCACGACAAAGGGGGTCTCTCCAACTCTGGAAGAGCTTAAAGA	6070
Qy	6237	AGAGCCACTTGAAGGGGAGCTGCGGCCCAAGACGACCGGCCCTGTGAAGCTTGAAGGAGG	6296
Db	6071	AGAGCCACTTGAAGGGGAGCTGCGGCCCAAGACGACCGGCCCTGTGAAGCTTGAAGGAGG	6130
Qy	6297	AGGCGGCCCACTCTCCACACTTGGGCGCGCTGCTGAGAGGCAAGCTCTGCTCAGCCCGC	6356
Db	6131	AGGCTGCGCCATCTCCCACTGTGGGCGCACTGCGCGAGGCAAGCCCTCATTCAGCCCAAC	6190
Qy	6357	TGCTTCAGACCGGCGCCGAGGGGTCAAAAGGTATCAAGCGGGGTGTCAACCTGGCCAGCA	6416
Db	6191	TCTCTCAAGACTGCGCCAGGACATCAAAAGGTATCAAGAGGTGTGTCACTTGGCTCAGCA	6250
Qy	6417	TCACTGAAGTATCAACAAGACTTAAACCCCGGACCAACCCACAGAGCTCAGCGCAACCC	6476
Db	6251	TCAAGAGAGTATTAAGCAGGACTTAACCCGGACCAACCCGACAGAGTCAAGTGAAGCCCGC	6310
Qy	6477	TGCGCGCGCCCTCTAATCTCTTCCCTGGAGGCAAGTGCCTCGTCTTGAACCTCGCGCGC	6536
Db	6311	TTCGCGCGCCCTCTAATCTCTTCCCGGAGGCAAGTGCCTCGTCTTGAATCTTGGCGCGC	6370
Qy	6537	CAGCCAGTGAATCTTAACTCCCGGCGCCGAGCAATGAGGCCCGGCGCGGTGAGCTCCCCC	6596
Db	6371	CAGCCAGTGAATCTTAACTCCCAACCCCGGACATGAGCAACCCAGACCCCGGGAATCCCCC	6430
Qy	6597	ACAGCGAAGGGGCAAGAGTCTCCAGAGCCAAACAAGACGTGCTTGGGTGTGTGTG	6656
Db	6431	ACAGTGAAGGGGCAAAAGTCTCCAGAACCCAGAAAAACATCGGTCTTGGGACGAGTG	6490
Qy	6657	AGGACGGTATTGAACCTGTGTCTCCACCGGAGGGCATGAGGAGCAGGGGCACTTCCGGA	6716

Oy	6491	AGATGACATTGAACCTGTGTGCCCAACAAGGGGCATATGACTGAACCGACATCTCGSA	6550
Oy	6717	GTCGCTGTGTACCCGCTGCTGTATCCGGGATGGGGAAACAGAAGAACCCACAGAGATATGGGCT	6776
Db	6551	GCGCTGTGTATCCCATCTCTGTATTTAGAAACGGGGAAACAGGGCGAGCC-- -CAGGATGGGCT	6607
Oy	6777	CCAAGTCTCCACGAGCAACCAACGACGCGCACGCTCTTTCAGGAAAGCTGACCCAGACA	6836
Db	6608	CTTAAGTCTCCAGGCAACACCAGCCGCGCACGCTCTTTCAGTAAAGTGTACTGAGACA	6667
Oy	6837	ACTCCGCGCATGTGTCAAGTCCAAAGNAGCAAGATCAACAAGAACTGAACAACCAACAAC	6896
Db	6668	ACTCCGCGCATGTGTAAAGTCCAAAGNAGCAAGATCAACAAGAAACTCAACAACCAACAAC	6727
Oy	6897	GGAATTAGCTGAATTAACAATACAGCAAGCTTGCGGACGAGATTTTCAATAATGCCCCCA	6956
Db	6728	GGAACGAGCCAGAAATTAACAATATGGCCAGCTTGCGGACGAGAAATTTTCAACAATGCCCCCA	6787
Oy	6957	TCAACCGGAACAAGGCTTATGACCAATATGAAGCGCAGGGGTGACGAAACATGCGACACA	7016
Db	6788	TCACTGGAAGCAGGCTTATGACCTGTAGAAAGCCAGGCGGTCAAGAACACGCCACACA	6847
Oy	7017	ACATGGGGCTGGAAGGCACTAATTTGAAGGCACTCATGGGTAATATGACCAAGTGGGAAG	7076
Db	6848	ACATGGGGCTTAGAGGCCATTTATGAAAAGCACTCATGGGTAATATGATCAAGTGGGAAG	6907
Oy	7077	AGTCCCGCGCTCAGGCGCAATGCTTTTAAACCTCTGAATGCCAGTGCACGCTGCC-	7135
Db	6908	AGCCCGCGCGCTCGGGGCCCAATGCTTTTAAACCTCTGAATGCCAGGCGCAGTCTGCCCG	6967
Oy	7136	--GCTGCTATGTCCCTAATACCGCTGTCTGACGGAACGGAATGACACAACACTCACTCGCCAG	7193
Db	6968	CTGCTGCTATGTCCCTAATACCACTGCTCTACGGAACGGAATGACACAACACTCACTCGCCAG	7027
Oy	7194	GTGGCGGGGGAAGAGCCAAAGGCTCTGTGSCAGAACCCAGACGCGCAAAGAACCAAGTCCCAG	7253
Db	7028	GTGGAAGTGGGAAAGCCAAAGGCTCTGTGSCAGAACCTTAGACGCGCAAAGAACCAAGTCCAG	7087
Oy	7254	CCCCGGGCTTGAGCACTGTGGGGAACCGAGCACCCCTGTGTCTCTCAAGTCACTCGAGGGAG	7313
Db	7088	CACCAAGGCTTAGCGTCCGAGAGACCGAACCCCTTGTGTCTCTCAAGTCACTCAAGAGGGAG	7147
Oy	7314	ACTGCAACCGCGGAGCGCGCTACCAACCGCGTGTGGGAGAACAGGCCCTCTGTCGCCAG	7373
Db	7148	ACTGCAATTCGCGCAACACACTCAACCAACGCTGTGTGGAGAACCGGCCCTCATCTCGAG	7207
Oy	7374	GTTCACAGGCATTTCCCTTAACAACCCCGTGATCATGCGGCTGACAGGCGGAGTGTATGCTT	7433
Db	7208	GTTCACAGGCATTTCCCTTAACAACCTTTGATTAAGAGCTTACAGGCAAGGTATCATGCGCT	7267
Oy	7434	CCCCACCCCAACGAGGCTCTCCCGCGGGGACAGCGGAGCCCTCTGCTGAGCCCCCAACGAGCT	7493
Db	7268	CCCCGCCCAACCTGTGCTTGTGGGCAAGGAGCGGAGCCCTTAGCTGTCTCCACACAGCGCT	7327
Oy	7494	GGAGCAGAGAGCCCAAGCTCACTGCTTGCTTGCACTGACGAGACACTTCCGACAGCGAG	7552
Db	7328	GGAGTGAAGAGCCCAAGCTCACTGCTTGCTTGCACTGACGAGACACTTCCGACAGCGAG	7386
RESULT 13			
US-10-087-192-650			
; Sequence 650, Application US/10087192			
; Publication No. US20020182586A1			
; GENERAL INFORMATION:			
; APPLICANT: Morris; David W.			
; APPLICANT: Engelhard, Eric K.			
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR			
; FILE REFERENCE: 529452000122			
; CURRENT APPLICATION NUMBER: US/10/087,192			
; CURRENT FILING DATE: 2002-03-01			
; PRIOR APPLICATION NUMBER: US 09/747,377			
; PRIOR FILING DATE: 2000-12-22			



2839 CCCGATGGCGACCCCGGGCCAAATGCTCAACCCAGAAAGCACTGAGACTGAGAGAGCT 2898  
1954 CCGGCTGGAGATCCCGGGCCAGTACTCGCCCGAGAAAGCCGCTGGAAGCTGAGAGAGCT 2013  
2899 GAAAGCAGCAGCGGCTGCCATTCGCCCATCGAGTCAACAAAGTCCATGAGCCCCCGG 2958  
2014 GAAAGCAGCAGCGGCTGCCATTCGCCCATCGAGTCAACAAAGTCCATGAGCCCCCGG 2070  
2959 GGAAGGACGAGCTCCCAACGAAGGACCTCCCGAGCCCAACCCGCAACCGGAAAACCTGCA 3018  
2071 GGAAGGACGAGTACCCCAAGCAGTTCCTGAGCTTCACCAACGAGAGCACTTACA 2130  
3019 GCGGAGAGAGCAGCGCCCTCAGCAGCTGGCAGCAGCCCGCGGAGAAAGAGAGAGAGCC 3078  
2131 GCGGAGAGAGTGAAGTGTCTCAGCAGTGGGAGAGAACTCCAGTGGCAAGTCCCGAGCC 2190  
3079 GGCACCCCGCGCGACAGGA-----GACCTTGGCAGCCGA 3114  
2191 AGTGCCTCGCGCGAAGAGGAGAGAAACCGGATTCCTTCCGCGCTTCCCACTGA 2250  
3115 GCGCCAGAGCTGCGCGGAGACCCCGCTTGGCTGAGACTTCGCGGCTGCCCTTCCCGGTCG 3174  
2251 GGGGCCAAAGCTACCGACTGAGCCCGCAGCTGGTATCGGCGCTGCCCTTCCCACTCC 2310  
3175 CCGCGGTGAGTGAATCAAGGCTCCCGGCAATGCGCGGAGCCCTCAGCTTCTCTTACG 3234  
2311 TCCAGGAGAGTGAATCAAGCTTCCCAACAGCGCGTGAACCTTCGTCTTCTCTTACAC 2370  
3235 TCCACTGTGATCAACCACTGCGCGCTGAGCTCAATGACATGCGCGCGCTGCTGCGCG 3294  
2371 ACCCGCGGATCAACCGCTGCTGAGGCTCCAGATGATGCGCGCGCTGCTGCGCG 2430  
3295 CCCACCCAGCTTCTCAACCCCGCTCCCTCATCTCTCTGCGCAACGCCAGCGCTCT 3354  
2431 ---TCCCCCATCTTCAACCCCCCAACCTCTCATCTCTGCGCAACCTCCCGCGTACT 2487  
3355 CGAGAGGCAATATAGTGCATCTCCCAAGAAATGTGAGTCCAGCTCAACGCTCCGACTC 3414  
2488 TGAAGAGCAGCTGAGTGCATCTCCAGAGGAGATGTAGTTCAGCTTCTGATGCTCACTC 2547  
3415 AGAGCATGCGCAAGGCCCGGTGGGCTCTGTCAACATGAGGCTGCGCTGCCATGAGACC 3474  
2548 AGAGCATGCGCAAGGCCCGGTGGGCTCTGTCAACATGAGGCTGCGCTGCCATGAGACC 2607  
3475 CAAAAAGCTGGCAACCTTCAAGCGGAGTGAAGAGAGAGAGCTGCCACGAGGCGCAAGG 3534  
2608 TAAAGAGCTG----- 2617  
3535 TGGGCGACGAGAGGCTGGGGGTGCCACAGCCGAGAGGCGTCCGTGTGAGAGGAGC 3594  
2618 -----GGGAC 2622  
3595 AGCTTGGGCTCAGTTCGCGGCGGAAGATCAACAAAGGATTCGAGCAACGAGGTGCC 3654  
2623 AGCAGTGGGCTCGGCGCACGATGGAAGATCAACAAAGGCGCTCCAGTACCGGG----- 2678  
3655 CTGCGACAGCGGCATCATACCGCGCTTCATCAACCCAGAGAGCGCGAGCTGACCTCT 3714  
2679 --CTGAGACGCGGCCGCTACAGAGGCTTATCAACCCAGAGCGCGCGCAACCTCTCT 2736  
3715 GTACAGAGGACCATCAACGAGATCATCGGCGAGAGCAGCGCGAGTCCGCTTGGAGCGCG 3774  
2737 CTACAGAGGATCATAGAGAGATCGTGGTGAAGACAGCCCAAGTCCGCTTGAACGGGCG 2796  
3775 CCGGAGAGCAGGCTGCCAAGGAGCGATCATCTTACGAAGGAGAGAGGCGCATCTT 3834  
2797 ACGAGAGAGCAGCTTCCCAAGGCGCATGTCACTTGAAGGAGCAAGAAAGCGCATCTCT 2856  
3835 GTCTCTATGAGGCTGATGTCTGTGACCAAGTCTCCAGAGAGAGAGGAGAGAGCTC 3894  
2857 ATCTCTATGAGGCTGATGTCTGTGACCAAGTCTCTTAAAGAGAGAGAGAGAGCTC 2916  
3895 AGGAGCCCCCATGAGAGGCGCGCCCAAGCGCATCTATGACATGATGAGAGGCGCGCT 3954

2917 GGGCCACCCCATGAGATGCGCGCCCTTAAACGCACTTATGACATGATGAGAGGCGCTGT 2976  
3955 GGGCAGAGCATATCTCTTACAGCAGATTCGAAGTCTATGAGGCGCTTCCCGCGCA 4014  
2977 AGGAGAGATGTTCATCTTACAGCAGATGAGAGGATCTATGAGGCGCTTCCCGCGCA 3036  
4015 GGCACAGAGCCCGCACCATCTTAAAGAGAGAGCAGCAGCATTCGCGGCTCATCAACAAG 4074  
3037 GC-----ACAGCCCGCATCTTAAAGAGAGATCACTCCAGGCTCATCAACGAGG 3090  
4075 GATCTCGGCTCTTACGTGAGAGCAGAGAGATCTAGCTGGAGAGCGCAAGCTCT 4134  
3091 CATCCGAGGTCTTATGTGAGGCGCAGAGAGCTATTACGCGGAGAGCGCAAGCTCT 3150  
4135 AAAGCGGAGAGGAGCGCTCCCGCCCATCCGCTTACCGGAGCTTGAACCGAGGCTTACAA 4194  
3151 GAAAGCAGAAAGGAGCAACACACCCCGACACCATTCGAGAGCTTGAAGCTTACAA 3210  
4195 GACGAGAGCG-----CTGGGCGCCCTGAAGCTGAAGCGGCGCATGAGGCGCTGT 4245  
3211 GCGCGGCGCGCTGAGACCTCTGAGTCCCTTGAAGTGAAGCGGACTCAAGAGGCTGTGT 3270  
4246 GGCACAGTGAAGAGAGCGGCGGCTCATTCATGAGATCCGCGGAGAGGCTGCGCG 4305  
3271 AGCACTGTAAGAGAGCGGCGGCTTATCATGATCCGAGAGAGAGGCTGCGCG 3330  
4306 CAGCGCGAGCTGCGCTTCCCGCGCGCTCAAGAGAGGCTCATCAACGAGGCGAC 4365  
3331 CACACTGAGCTACCCCTGGCACACGCGCTCTGAAGAGGCTTCATCAACCGAGGCGAC 3390  
4366 CCGGCTCAAGTACGACACCGGCGGCTTCACTGAGCTCAAAAGACAGACGTAAGCTC 4425  
3391 CCACTTCAAGTACGACTTGGGAGACCTTCATGAGCAGAGAAACAGACGCTGCGCTC 3450  
4426 CTTGATCGCAGCGCCCGCGCGAGTTCACACCGCTGCAACCGCTGATGTGATGCCGA 4485  
3451 CATCATCGCAGCGCGCGCGCTTCCCTGCGCTGCAACCGCTGACATTAATGCTGA 3510  
4486 CCGCGGCGCATGGAACGTGCTGCTACAGAGAGCTGGAAGAGCGCGCGCAGGAGCCG 4545  
3511 TGCCTGGGCACTGAGCGCTGCTATGAAGAGTCTGAAGAGCGGCTCAGGAGCACG 3570  
4546 CAGAGCTCGGGGGGCTCATTTGGCGGCGCGCGCTCATTTGAGCTGAGGCTGA 4605  
3571 CAGTGTGAGAGGAGGCTCATACAGTGGGCTTCAAGTGTGCTGAACTTGGGCA 3630  
4606 GCGCGGAGAGCGCCCTGAGCTTATGAGAGCAGCGGAGCACCTTGGCGCGCATCTCC 4665  
3631 GCGAGGCAAGGCGCATGATTTAGAGAGACAGCGGAGCACCTTCAACAGTCACTGCC 3690  
4666 ACGAGTTTCCCGCTGACATGCGGAGCGCACGCGCGCTGACAGAGGAGCGCTTTC 4725  
3691 ACGTGTCTCCCTGTGACACAGAGGAGCGCACGCGCATCTTCAAGAAAGCGCTCTCT 3750  
4726 GTTCAGCAAGGAGATCCAGAGCCGAAGGCTGATGAGCGCTGATGAGTGGCAAGT 4785  
3751 ATTCAGCAAGGAGTCCAGAGCCGAAGGCTGATTTAACCCCGGAGATGGCAAATC 3810  
4786 CCGCAGACAGCCTGCGCGGAGCACACCAACCCCATCTGCGCTTATGAGCACTGCT 4845  
3811 CCGCAGACAGCCTGCGCGGAGCACACCAACCCCATCTGCGCTTATGAGCACTGCT 3870  
4846 TGGGCGGTGATGAGCGCTGAGCTTATGAGCAGCAGCATTCCTTGGCTTGAACCCAC 4905  
3871 CCGGCGGTGATGAGCGCTGAGCTTATGAGCAGCAGCATTCCTTGGCTTGAACCCAC 3930  
4906 CTTCATACCCCGGAGATCCCTGTGA-----CGAGCGCTGACTTACCTGCGCGGACA 4962  
3931 CTTCATACCCCGGAGATCCCTGTGAAGAGCAGCAGCGCTTACCTGCGCGGACA 3990  
4963 CTTGGCGCGCAACCCGACTTACCGGACTTGAACCACTTACCTGATCGCGGCTAACCC 5022

Db 3991 CTTGGCCCCAGCCCCCACTTACCACCTGTATCCCACTTACCTTCATCCGCGCTACCC 4050  
Qy 5023 CGACA CGCGCGGCTGTAGAGAACCGGAGACCAATCATTAATGAATGATCATCATCTCGAGCA 5082  
Db 4051 TGACACGGCGGCGCTGTAGAGAACCGGAGACCAATCATTAATGAATGATCATCATCTCGAGCA 4110  
Qy 5083 GATGACCAACAACAGGCGCACTGGCCATGGCCCAAGGAGCTGATATGCTGAGGGGCTCTTC 5142  
Db 4111 GATGACCAACAACAGTGTCTCGCCATGGCCCAAGGAGCTGATGATGAGGGGCTCTTC 4170  
Qy 5143 GCCCCGAGTCTGTGGTGGCACTAACATAGCTGGGGTCCCGAGGATGATGAGACT 5202  
Db 4171 ACCGGAGAGTCTGTGGTGGCCCTTAATTAATGCGCTGGCCCAAGAGGATTAATGACT 4230  
Qy 5203 GTTCCAAAGTGCACACTGCTGTGTCTGTGCCCCGACACCAAGGACCCCAAGCAACGCG 5262  
Db 4231 GTTCCAAAGTGCACACTGCGGTGTGTGCAACAAGGACGAGGCCCTCGACACGCG 4290  
Qy 5283 CATGACCGCTTGGCTTACTTCCCAACGCGGCCCAAGCCCTTACAGAGCTGCAACAGAG 5322  
Db 4291 CATGACCGCTTGGCTTACTTCCCACTGCGGCCCAAGCCCTTACAGAGCTGCAACAGAG 4350  
Qy 5323 CTCCCACTCTCTCCAGAGAGTCAACAACATTAACAAACCAACCAAGCTCTCGT 5382  
Db 4351 CTCACTGCTGTCCAGAGAGCCCACTTACCTTAAGTAAACCACTGCAATCTTCATC 4410  
Qy 5383 CGAGCGGAGCGAGACCGGGATCGAGAGCGGAGACCGGATCGGAGCGGAGAAAGTCAAT 5442  
Db 4411 GAGAGCG-----GAAAGGAAAGTGAAGCGGAAACAGAGACAATCTCAT 4452  
Qy 5443 CTCACTGCTTACACAGACGCTGAGACAGCAACCAATCTGAGAGACTGTGTAACAGAGAG 5502  
Db 4453 CTCACTGCTTACACATGATGAGACATGCAACCACTGAGAGACTGTGTAACAGAGAG 4512  
Qy 5503 CAGCGGAGCAAGCGGAGAGCGAGCGGGGATGGGGGAGAGCAACCGCGCCCTCCCA 5562  
Db 4513 CAGCGGGCTGGGGAGAGAGCG-----CCCGCG 4542  
Qy 5563 CTCCCATGCCACAGCACTCGCCCATCTCCCTCGAGACCAAGATGCTTCAGACAG 5622  
Db 4543 CTCCCAACCCACAGCACTGGCCCATCTCCCGGAGACCAAGAGCTTGTGACACAGAG 4602  
Qy 5623 ACCGAGTGTCTTCAACAACAGGCAATGAAGGTATCATCACG-GCTGTGAGCCCAAG 5681  
Db 4603 GCCAGTGTGTGCAACAACAGGCAATGAAGGCTGTGTCACTCCCGTGAACCCGAG 4662  
Qy 5682 AGCCACGAGTCTGAGTGTCACTGCACTCTCAACCGTTCGCCAGTGTGCAATTC 5741  
Db 4663 CGCCACGAGTCTGAGTGTCACTGCACTCTCTTGGCTGTCCGACGCTGCAATTC 4722  
Qy 5742 CACTTGCAACCACTGCACTGGAGCGGCACTCTGATGGAGTCTTACCTCATAG 5801  
Db 4723 CACTTGCAACCACTGCACTGGAGCGGCACTTGAAGGGTCTTACCTTACCTCATAG 4782  
Qy 5802 AGCCGTCTTGTGTGCAAGAGAGCGCCCGGGTGTCCCGGACAGAGCGCCGAGAG 5861  
Db 4783 AGCCGTCTTGTGTGCAAGAGAGCTGTCCGGTGTCCCGGACGAGGCGCCGCTGTAG 4842  
Qy 5862 ACAACGAGCATGCTCTCTGCAAGCCCCCAAGCCCGGCTCGCGGGTGTGAGCCCGCTCT 5921  
Db 4843 AGCGTGGCATGCTCTCTCTCAACAAACCCCGGCG-----GAGGCCGCTCTCT 4893  
Qy 5922 CCCCCAGAAAGGAGTGTGAGGCCCCCGGCTTATGATGCTCTGTCTGTGCAACGACCA 5981  
Db 4894 CACCGAGAGAGGCTCGAGGCCCGCATTCCTTAGACACCCCGAGCTTCAGACCAACAGCCA 4953  
Qy 5982 TCGCCCGCAACCCCTGTGAAACCTTGCACTTCACAGCGCAAGCCGAGACCCGCGGCGC 6041  
Db 4954 TCGCCCGCAACCCAGCAAAAGAGCTTTGACCCCAACATGCAAGTCCGAGACCGCGGCG- 5012  
Qy 6042 CACGTGCTCGGCTCGAGACCCGAGACCGGAGAAACATCAAAAGTAAACCTTTTCATTC 6101  
Db 5013 --CCACCTCGGCTCAAGATCTGACCGAGAAAGATCAAAAGTAAACCTTTTCATTC 5070

Qy 6102 AGGAATGAACTCCGTTCTGTGGTTACAC---GGAGAGCTTACAGCCCCGAGGG 6158  
Db 5071 AGGAATGAACTCCGTTCTGTGGTTACACAGTGGAGGTGGTTACAGCCCCGAGTGGG 5130  
Qy 6159 TGGAGCCCGTCAAGCCCTGTGAGCTACCCAGTGTGACCCAGCAAGAGGGGCTCCCAAG 6218  
Db 5131 TGGAGCCCATCAAGCCCGTGTGAGCTCCCAAGCTTACCAACAGAGGGGCTCTCAAC 5190  
Qy 6219 ACCTGAAGAGCTCGACAAGAGCCACTGTGAGGGGAGGTGGCGGCCAAGAGCCAGGCC 6278  
Db 5191 CTCTGAAGAGCTTGAAGAGAGCCACTTGAAGGGAGTGTGGGACAGACAGCCAGGCC 5250  
Qy 6279 CCGTGAAGCTTGGCGGAGAGGCGGCCCACTTCCCAACCTGTGCGCTGTCTGAGAGCC 6338  
Db 5251 CCATGAGCTCAGCGCGAGGCTGCCATTCCTCCACATCTGGGCGCACTGCCGAGAGCC 5310  
Qy 6339 AGCCCTCGTCCAGCCCGCTGTCCAGACCGGCTCCAGGGGTCAAAGTCAACAGCGGGTGG 6398  
Db 5311 AGCCCTCATCCAGCCCACTCTCCAGACTGCCCCAGGATCAAAAGTCAACAGAGGTGG 5370  
Qy 6399 TCACCTGAGCCAGACATCAGTGAATCATCAACAGAGATCAACCCGAGCAACCCAC 6458  
Db 5371 TCACCTGAGCTCAGACATCAGAGAGTCAATTAGAGAGATCAACAGCGGCCACCCGC 5430  
Qy 6459 AGCAGCTAGGCGCAACCTGCGCGCCCTCTTACTCTCTTCCGTGGGGCAGCTGCCG 6518  
Db 5431 AGCAGCTAGTGGCGCCCTTCCCGCCCTCTCTACTCTCTTCCGGAGCGAGCTGCCG 5490  
Qy 6519 TCTTGAAGCTCGCGGCCCACTGATGACTGTACTCTCCGCGCCCGGACCAATGATGCC 6578  
Db 5491 TCTTGAATCTTCGCGGCCCACTGATGACTGTACTCTCCACCCCGGACCAATGACACC 5550  
Qy 6579 CGGCGCGTGGCTCCCGCCACAGCGGAGGGGCAAGAGGTCTCAGAGCAACAGAGT 6638  
Db 5551 CAGCCGGAGATTCGCCCAAGTGAAGGGGCAAAAGTCCCAAGACCCAGCAAAAT 5610  
Qy 6639 CGGTCTTGGAGTGTGTGAGAGCGGTATTGAACCTGTGTCCCAAGGAGGATGAGCG 6698  
Db 5611 CGGTCTTGGAGGAGGATGAGAGTGCATTGAGCGTGTCCCAACAGAGGCAATGAGCG 5670  
Qy 6689 AGCCAGGCACTCCCGGAGTGTGTGACCGGTGTGTACCGGAGTGGGAAACAGAGCG 6758  
Db 5671 AGCCAGGCAATGCTCGAGGCGCTGTGTACCACTGTATTCAGAACGGGAAACGAGCG 5730  
Qy 6759 AGCCAGGAGATGGGGCTCAAGTCTCCAGGCAACAGAGCGGCAAGCTTCTTCA 6818  
Db 5731 AGCC---CAGATGGGCTTAAATCTTCCAGGCAACAGAGCGGCAAGCTTCTTCA 5787  
Qy 6819 GCAAGCTGACGAGAGCACTCCGCAATGTCAAGTCCAGAGCAAGAGCAAGATCAACAGA 6878  
Db 5788 GTAAGCTGACGAGAGCACTCCGCAATGTGAAGTGAAGAGAGAGATCAACAGA 5847  
Qy 6879 AGCTGAACCAACCAACCGGAATGACTGAATACATATCAAGCCAGCTGGAGCGAGA 6938  
Db 5848 AACTCAACCAACCAACCGGAACGAGCAAGATCAAAATATGGCCAGCTGGAGCGAGAA 5907  
Qy 6939 TCTTCAATATGCGCGCATCAACGGAACAGGCTTATGACCTATAGAACAGAGGCGAGC 6998  
Db 5908 TCTTCAATATGCGCGCATCACTGAGAGAGGCTTATGACCTGTAAGAACAGAGGCGAGC 5967  
Qy 6999 AGGAACATGCAAGCAACAATGGGGCTGAGAGCATTAATAGAAAGCACTCATAGGTA 7058  
Db 5968 AAGGAACAGCAAGCAACAATGGGGCTGAGAGCATTAATAGAAAGCACTCATAGGTA 6027  
Qy 7059 AATATGACAGTGGAGAGTTCCTCGCGCTCAGCGCAATGCTTTTAACTCTGAATG 7118  
Db 6028 AATATGATCAATGGAGAGAGCCCGCGCTCGGCGCAATGCTTTTAACTCTGAATG 6087  
Qy 7119 CCAATGCAAGCTGCC---GCTGTATGCGCATTAACCGCTGTGACGAGAGAGTGAAC 7175  
Db 6088 CCAAGGCAAGTGTGCGCGCTGTGTATGCCCATTAACCACTGTGACGAGAGAGTGAAC 6147

QY 7176 ACACACTCACTTCGCGAGGTGGCGCGGAGAGCCAAAGTCTCTGACAGACCCAGCAGCC 7235  
 DB 6148 AGGCACTCACTTCGCGAGGTGGAGGTGGAGAAAGCCAAAGTCTCTGACAGACCCAGCAGCC 6207  
 QY 7236 GAAAAGCCAAAGTCCCGGCGCGGCGCTTGGCATCTGGGAGACCGGCAACCTCTGTCTCTCT 7295  
 DB 6208 GAAAAGCCAAAGTCCCGGCGCGGCGCTTGGCATCTGGGAGACCGGCAACCTCTGTCTCTCT 6267  
 QY 7296 CAGTGCACTCGAGGAGGAGCTGCAACCGGCGGAGCGCGCTGCAACCGGCTGTGGAGG 7355  
 DB 6268 CAGTACACTCAAGAGGAGGAGCTGCAATCGCGAACAACCACTCAACCAACCGTGTGGAGG 6327  
 QY 7356 ACAGGCGCTCTGTCGCGAGGTTCACGCGCATTCCTCTCAACACCTCTGATCATGCGGCTCG 7415  
 DB 6328 ACCGCGCTCTGATCTGAGGGTTCAGCGCATTCCTCTCAACACCTCTGATTAAGGCTAC 6387  
 QY 7416 AGGCGGGTGTCAATGCTTCCCAACCCCAACCGGCGCTTCCCGCGGAGCGGCGCTCTCG 7475  
 DB 6388 AGGCAAGTGTCAATGCTTCCCGCGCTTCCCGCGGAGCGGCGGCGCTCTAG 6447  
 QY 7476 CTGGCCCCCAACCGCTGGGAGCAGAGAGCCCAAGCCACTGCTGTGCTGCGAGTACAGAGA 7535  
 DB 6448 CTGGTCTCCCAACCGCTGGAGTGAAGAGCCCAAGCCACTGCTGTGCTGCGAGTATGAGA 6507  
 QY 7536 CACTCTCCGACAGCGAGTACTCAGAACAGGCGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCT 7595  
 DB 6508 CACTCTCCGACAGCGAGTACTCAGAACAGGCGGCGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCT 6549  
 QY 7596 GTCAAGTCTCCAGCGAGCCACAGAGAACGCTCTGACAGAGCGGCGGCGGCGGCGGCTCTCC 7655  
 DB 6550 GCCAGGTCTCCCGCACAAGGAGAGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6609  
 QY 7656 CAACCAAGAGAGAGGCGCTGAGTCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7699  
 DB 6610 AGACTGAGAGAGAGGCGCTGAGTCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6669  
 QY 7700 CTGCTCGTCCAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7759  
 DB 6670 CTGCTCATCTCAAGAGTGGGATCC-TCCTGTCTTAAGGCTTAAAGTCTCAAGGCTCCACCGG 6728  
 QY 7760 GCGTGGCGCTGTGACG---ACCTTACTCAAGGAGTGTAACTGCTGTGCTGTGAGAGGAG 7815  
 DB 6729 GCGTGGCGCTGTGACGAGTGAACCTTAACCTCAAGGAGTGTAACTGCTGTGCTGTGAGAGGAG 6788  
 QY 7816 GGGGA-----AGGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7858  
 DB 6789 GTGAGCAGAGGAGGAGGAGCAAGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6848  
 QY 7859 ACACACAGGCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7918  
 DB 6849 GCACTCTGTGATCTGTGGGAGGAG-CGGGAGACCCACACAGGCTGACAGGCACTTCCATGCGC 6907  
 QY 7919 ACTGCT--CCCCGAGATCTTTGGAACCAAGTCTAACTGAGCTGCAAGCCCGCGG 7976  
 DB 6908 ACCGCTGCGCTTAAACGCAATTTGGAACCAAGTCTAACTGAGCTGCAAGCTGCGGTGCTCT 6967  
 QY 7977 CCTCTCTGCGGCTCCGATCCGCTTAAAGGCTGTGAGCAGATGAGAGGAGGCGGCTGTCCA 8036  
 DB 6968 GCGCTCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTATTC 7027  
 QY 8037 GCGCGCGAGTGGCGTGTGCGGAGTCCCAAGACTGCGCGCAAGCGAAGAGATTTGTGGAGA 8096  
 DB 7028 CTGCGCTTAATGCTGTATCATGAGCATCTTCAAGGCGCAACCGCAACAGACTTGGAG 7087  
 QY 8097 CCAAGTCAAGCGAGTGGCGGAGCAAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 8156  
 DB 7088 CCA-TCAGACCAAGGAGGAGGAGC-AAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7143  
 QY 8157 CCGAGGAGTCAAGTGTTTTTCACACATCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 8216  
 DB 7144 TGCAGGAGTCAAGCCGTTTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7199  
 QY 8217 TGTAAATGATGTGTGTTTACAGGAGTATATTTTGTATACCTTCAATGATTAATTGAGA 8276

DB 7200 TGTAAATG--GTGTGGCTTACAGGAGTATATTTTGTATACCTTCAATGATTAATTGAGA 7257  
 QY 8277 TGTATTAACCAAGAGAGGAGCTTAACCAAGTATACGCTGCTGCTGCTTGTATCTGCTT 8336  
 DB 7258 CGTCTCAACAGAGAGAGAGCTGCGCC-GTGTTCGTGCGGCTGTGC-TTGTGTCTTACCT 7315  
 QY 8337 ACCGTTCAAGAGGCGGCTGACGAGCGGAGCAGTGGTGAACCCCATCACTCCAGAGACCAAG 8396  
 DB 7316 ACTGTTTCAAGAGGAGAGCTGCGAGCC--AAGCGGTGGCGGCGGCGGCGGCGGCGGCGGCGG 7373  
 QY 8397 GGGCGGAGAGTGTCTGTCAAGCGCGGCTG-TGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8455  
 DB 7374 GGTCAAGGAGGCTCCGAGCAGCGGCACTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7433  
 QY 8456 ATGAAATCAATGCGATTTCTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 8515  
 DB 7434 ATGGAATGATGCGATTTCTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7490  
 QY 8516 ACACAGTGTCTTAATTAAGGCAAT 8544  
 DB 7491 ACACATGTTTCTCAATTAAGGCAAT 7519

RESULT 14  
 US-10-146-473-14  
 ; Sequence 14, Application US/10146473  
 ; Publication No. US2003010888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Scanlan, Matthew  
 ; APPLICANT: Gout, Ivan  
 ; APPLICANT: Stockert, Elisabeth  
 ; APPLICANT: Gure, Ali  
 ; APPLICANT: Chen, Yao-Tseng  
 ; APPLICANT: Old, Lloyd  
 ; TITLE OF INVENTION: Breast Cancer Antigens  
 ; FILE REFERENCE: L00461/70130 (JRV)  
 ; CURRENT APPLICATION NUMBER: US/10/146,473  
 ; PRIOR FILING DATE: 2002-05-15  
 ; PRIOR APPLICATION NUMBER: US 60/291,150  
 ; NUMBER OF SEQ ID NOS: 82  
 ; SOFTWARE: Patentin version 3.0  
 ; SEQ ID NO 14  
 ; LENGTH: 2930  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-146-473-14

Query Match 30.6%; Score 2618.6; DB 5; Length 2930;  
 Best Local Similarity 94.6%; Pred. No. 0;  
 Matches 2786; Conservative 3; Mismatches 9; Indels 146; Gaps 2;  
 QY 4785 CCCCCGACAGACCGTGCCTGCGGAGCAGACCCACATCTGCGCTTATGAGACTGCG 4844  
 DB 133 CCCCCGACAGACCGTGCCTGCGGAGCAGACCCACATCTGCGCTTATGAGACTGCG 192  
 QY 4845 TTGCGGCGCTGAGTGGCGTGAACCTGTATGCAAGCCACATCCCTCTGAGCTTGAACCCA 4904  
 DB 193 TTGCGGCGCTGAGTGGCGTGAACCTGTATGCAAGCCACATCCCTCTGAGCTTGAACCCA 252  
 QY 4905 CCTTCATACCGCGCGGAGTCCCTCTGAGAGCAGCGGCTGCTTACTTACCTGCGCGGAGC 4964  
 DB 253 CCTTCATACCGCGCGGAGTCCCTCTGAGAGCAGCGGCTGCTTACTTACCTGCGCGGAGC 312  
 QY 4965 TGGCCCCCAACCCCACTTACCGGCACTGTACCAACCTTACTTATCCGCGGCTACCCG 5024  
 DB 313 TGGCCCCCAACCCCACTTACCGGCACTGTACCAACCTTACTTATCCGCGGCTACCCG 372  
 QY 5025 ACACGCGGCGCTGAGAGAACCGGAGACCATATCAATGATCAATCACTTGGCAGAGAA 5084  
 DB 373 ACACGCGGCGCTGAGAGAACCGGAGACCATATCAATGATCAATCACTTGGCAGAGAA 432

QY 5085 TGCAACAACAACGCGCACCCGCAATGAGCCGCAAGCTGATATGCTGAGAGGCGCTCTCGC 5144  
DB 433 TGCAACAACAACGCGCGCACCCGCAATGAGCCGCAAGCTGATATGCTGAGAGGCGCTCTCGC 492  
QY 5145 CCGCGAGTCTCCGCGTGGCACTCACTAGCTGCGGGTCCCCGAGGCAATCACTGACCTGT 5204  
DB 493 CCGCGAGTCTCCGCGTGGCACTCACTAGCTGCGGGTCCCCGAGGCAATCACTGACCTGT 552  
QY 5205 CCGAAGTGCACAACCTGCGCTGTGCTGTGCCCCGACACAGAGCAACCCAGCGCACCGCA 5264  
DB 553 CCGAAGTGCACAACCTGCGCTGTGCTGTGCCCCGACACAGAGCAACCCAGCGCACCGCA 612  
QY 5265 TGGACCGCTCTGCTCACTCTCCCAACGCGGCCGAGCGCTTTCAGACGCGCGCACAGAGCT 5324  
DB 613 TGGACCGCTCTGCTCACTCTCCCAACGCGGCCGAGCGCTTTCAGACGCGCGCACAGAGCT 672  
QY 5335 CCGCACTCTCCCGAAGAGTCCACAACAATTGACAAAACCAACCAACGATCTCTGTCCG 5384  
DB 673 CCGCACTCTCCCGAAGAGTCCACAACAATTGACAAAACCAACCAACGATCTCTGTCCG 732  
QY 5385 AGCGGAGCGAGAACGCGGATCGAGACGCGGACCGGGATCGGAGCGGGAAGTCCATCC 5444  
DB 733 AGCGGAGCGAGAACCGGGATCGAGACGCGGATCGGAGCGGGAAGTCCATCC 792  
QY 5445 TCAAGTCCACAACAACGCGTGAAGCAACGATCTGAGACCTGATACAGAGCAGACA 5504  
DB 793 TCAAGTCCACAACAACGCGTGAAGCAACGATCTGAGACCTGATACAGAGCAGACA 852  
QY 5505 GCGGAGAGAGCGGCGACAGCGCGCGGGGTGGGGGAGAGCAGACGCGCGCGCGCTCCACT 5564  
DB 853 GCGGAGAGAGCGGCGACAGCGCGCGGGGTGGGGGAGAGCAGACGCGCGCGCGCTCCACT 912  
QY 5565 CCGCATGCGCACACAGACTCGCCCATCTCCCTCGAGACCAAGGATGCGCTCCAGCAGAGAC 5624  
DB 913 CCGCATGCGCACACAGACTCGCCCATCTCCCTCGAGACCAAGGATGCGCTCCAGCAGAGAC 972  
QY 5625 CCAAGTGTCTTCAACAACAAGGATGAAGGATATCAACGCTGTGAAGCCGACGAAGC 5684  
DB 973 CCAAGTGTCTTCAACAACAAGGATGAAGGATATCAACGCTGTGAAGCCGACGAAGC 1032  
QY 5685 CCAAGTGTCTTCAACAACAAGGATGAAGGATATCAACGCTGTGAAGCCGACGAAGC 5744  
DB 1033 CCAAGTGTCTTCAACAACAAGGATGAAGGATATCAACGCTGTGAAGCCGACGAAGC 1092  
QY 5745 CTGCAACCCACATGCGCGCACCTGAGTGGGGTCTAAGCTTACCTCATGAGAC 5804  
DB 1093 CTGCAACCCACATGCGCGCACCTGAGTGGGGTCTAAGCTTACCTCATGAGAC 1152  
QY 5805 CCGTCTTGTCTGCGCAAGAGGCGCCCGGGGTGCGCGCGCAGAGCGCGCCGCGAGACA 5864  
DB 1153 CCGTCTTGTCTGCGCAAGAGGCGCCCGGGGTGCGCGCGCAGAGCGCGCCGCGAGACA 1212  
QY 5865 CCGGCGCATGCGCTTCTGCGCAAGCGCCCGGCTCGGGGCTGGAAGCCGCGCTCTCC 5924  
DB 1213 CCGGCGCATGCGCTTCTGCGCAAGCGCCCGGCTCGGGGCTGGAAGCGCGCTCTCC 1272  
QY 5925 CCAAGCAAGGCTCGAGAGCGCGCGCGCGCTAAGTGCCTGTCTGTGCGACACGACCATG 5984  
DB 1273 CCAAGCAAGGCTCGAGAGCGCGCGCGCGCTAAGTGCCTGTCTGTGCGACACGACCATG 1332  
QY 5985 CCGCGACCCCTGCGAAGAACTTGACCTCAACAGCGCAAGCCCGAGCCGCGGCGCAC 6044  
DB 1333 CCGCGACCCCTGCGAAGAACTTGACCTCAACAGCGCAAGCCCGAGCCGCGGCGCAC 1392  
QY 6045 CTGCGCTCGGCGCTCGAGACCGCGCAACCGGGAAGAACTCAAGTAAACCTTTTTCATCCAG 6104  
DB 1393 CTGCGCTCGGCGCTCGAGACCGCGCAACCGGGAAGAACTCAAGTAAACCTTTTTCATCCAG 1452  
QY 6105 AACTGGAATCTCGGTTCTGAGGTACCAAGCAGAGCTACAGCCCGAAGGGGTGAGGC 6164  
DB 1453 AACTGGAATCTCGGTTCTGAGGTACCAAGCAGAGCTACAGCCCGAAGGGGTGAGGC 1512  
QY 6165 CCGTCAAGCCCTGTGAGCTCAACCAAGTCTGACCAAGCAAGGGGCTCCCAAGCACTGG 6224

DB 1513 CCGTCAAGCCCTGTGAGCTACCCAGTCTGACCCACGAACAAGGGGCTCCCAAGCACTGG 1572  
QY 6225 AAGAGCTGACAAGAGCCACCGTGAAGGGGAGCTGCGGCCCAAGCAGCGAGCCCGTGA 6284  
DB 1573 AAGAGCTGACAAGAGCCACCGTGAAGGGGAGCTGCGGCCCAAGCAGCGAGCCCGTGA 1632  
QY 6285 AGCTTGGCGGGAGAGCGCGCACCTGCCACAACCTGCGCGCGCTGCTTGAAGACCACT 6344  
DB 1633 AGCTTGGCGGGAGAGCGCGCACCTGCCACAACCTGCGCGCGCTGCTTGAAGACCACT 1692  
QY 6345 GGTCCAGCCCGCTGTCCAGACCGCGCCCAAGGGGTAAAGTTCACAGCGGGTGTCAACC 6404  
DB 1693 GGTCCAGCCCGCTGTCCAGACCGCGCCCAAGGGGTAAAGTTCACAGCGGGTGTCAACC 1752  
QY 6405 TGGCCACAGCATCATGATGAGGTATCAACAAGACTACACCGGCAACCAACGACGAGC 6464  
DB 1753 TGGCCACAGCATCATGATGAGGTATCAACAAGACTACACCGGCAACCAACGAGCAGC 1812  
QY 6465 TCAAGCAACCCCTGCGCGCCCGCTCTACTCTCTCTGAGGCGAGCTGCCGTCTGG 6524  
DB 1813 TCAAGCAACCCCTGCGCGCCCGCTCTACTCTCTCTGAGGCGAGCTGCCGTCTGG 1872  
QY 6525 AACTCCGCGCGCACCCAGTGAACCTTACCTTCGCGCCCGGACCAATGTGCCCCGCGCC 6584  
DB 1873 AACTCCGCGCGCACCCAGTGAACCTTACCTTCGCGCCCGGACCAATGTGCCCCGCGCC 1932  
QY 6585 GTGGCTCCCCCAACGAGAGGGGCAAGAGGTCTCCAGAGCAACAAGAGTGGCT 6644  
DB 1933 GTGGCTCCCCCAACGAGAGGGGCAAGAGGTCTCCAGAGCAACAAGAGTGGCT 1992  
QY 6645 TGGGTGTGTGAGGACGGTATTGAACCTGTGTCCACCGAGGGGCAATGACGAGCAG 6704  
DB 1993 TGGGTGTGTGAGGACGGTATTGAACCTGTGTCCACCGAGGGGCAATGACGAGCAG 2052  
QY 6705 GGCACCTCCCGAGTGTGTGTACCGGCTGTGTACCGGGATGGGGAAACAAGACGAGCCCA 6764  
DB 2053 GGCACCTCCCGAGTGTGTGTGTACCGGCTGTGTACCGGGATGGGGAAACAAGACGAGCCCA 2112  
QY 6765 GCAAGATGGGCTCCAAAGCTCCAGGCAACACAGCAGCGCGCAGCTTCTTCAAGCAGC 6824  
DB 2113 GCAAGATGGGCTCCAAAGCTCCAGGCAACACAGCAGCGCGCAGCTTCTTCAAGCAGC 2172  
QY 6825 TGAACGAGAGCAACTCCGCAATGCTCAAGTCCAAAGCAAGAGATCAACAAGAGCTGA 6884  
DB 2173 TGAACGAGAGCAACTCCGCAATGCTCAAGTCCAAAGCAAGAGATCAACAAGAGCTGA 2232  
QY 6885 ACAACCAACAACCGGAATGAGCTGAAATATCAAGCAAGCTGCGACGAGATCTTCA 6944  
DB 2233 ACAACCAACAACCGGAATGAGCTGAAATATCAAGCAAGCTGCGACGAGATCTTCA 2292  
QY 6945 ATATGCCCCGCATCAACCGGAACAGGCGCTTATGACCTATGAAGGCAAGCGGTGAGAGC 7004  
DB 2293 ATATGCCCCGCATCAACCGGAACAGGCGCTTATGACCTATGAAGGCAAGCGGTGAGAGC 2352  
QY 7005 ATGCAAGCAACAATGAGGGCTGAGAGCCATTAATTGAAGGCACTCATGGTAAATATG 7064  
DB 2353 ATGCAAGCAACAATGAGGGCTGAGAGCCATTAATTGAAGGCACTCATG----- 2402  
QY 7065 ACCAGTGGGAAGAGTCCCGCGCTCAGCGCAATGCTTTTAACCTCTGAATGCCAGTG 7124  
DB 2403 ----- 2402  
QY 7125 CCAAGCTGCGCGCTGTGATGCGCATTAACCGGTGTGAACGAGGTGACCAACACTCA 7184  
DB 2403 ----- 2402  
QY 7185 CCGTGCAGATGTGCGCGGAGAGGCAAGGTCTGTGCGAGACCCAGCAGCGGAAAAAGCCA 7244  
DB 2493 -----GTTGGCGCGCGGAAAGGCAAGGTCTGTGCGAGACCCAGCGGAAAAAGCCA 2454  
QY 7245 AGTCCCGGCGCCGCGGCTGAGATGTGGGAACCGGCAACCTTGTCTTCAAGTCACT 7304





QY 6045 CTGCTCGGCTCGGACCCGCAACGGAAAGACTCAAAAGTAAACCTTTTTCATCCAGG 6104  
Db 1393 CTGCTCGGCTCGGACCCGCAACGGAAAGACTCAAAAGTAAACCTTTTTCATCCAGG 1452  
QY 6105 AACTGGAATCTCGTCTCTGAGGTTTACACGCGAGAGACTACAGCCCGAAGGGGGTGGAGC 6164  
Db 1453 AACTGGAATCTCGTCTCTGAGGTTTACACGCGAGAGACTACAGCCCGAAGGGGGTGGAGC 1512  
QY 6165 CCGTCAGGCTTGTAAGTCAACCAAGTCTGACCAAGACAAAGGGGCTCCCAAGCACTGG 6224  
Db 1513 CCGTCAGGCTTGTAAGTCAACCAAGTCTGACCAAGACAAAGGGGCTCCCAAGCACTGG 1572  
QY 6225 AAGAGCTGACAGAGCCACTCTGAGGGGAGCTGCGGCCAAGACCAAGGCCCTGGTA 6284  
Db 1573 AAGAGCTGACAGAGCCACTCTGAGGGGAGCTGCGGCCAAGACCAAGGCCCTGGTA 1632  
QY 6285 AGCTTGGGGGGAGGCGCGCCCACTCCCAACCTCGCGCGCTGCTGAGAGCCAGCCCT 6344  
Db 1633 AGCTTGGGGGGAGGCGCGCCCACTCCCAACCTCGCGCGCTGCTGAGAGCCAGCCCT 1692  
QY 6345 CGTCCAGGCGCGCTGCTCAGACCGCGCCGAGGGTCAAAAGTCAACGAGGGGTGTCACCC 6404  
Db 1693 CGTCCAGGCGCGCTGCTCAGACCGCGCCGAGGGTCAAAAGTCAACGAGGGGTGTCACCC 1752  
QY 6405 TGGCCCAAGACATCTAGTGAAGTCAATCAACAGACTCAACCCGCAACCAACAGCAGC 6464  
Db 1753 TGGCCCAAGACATCTAGTGAAGTCAATCAACAGACTCAACCCGCAACCAACAGCAGC 1812  
QY 6465 TCAGGCAACCCCTGCGCGCGCTCTACTCTCTTCCCTGGGGCGAGCTGCGCGCTGG 6524  
Db 1813 TCAGGCAACCCCTGCGCGCGCTCTACTCTCTTCCCTGGGGCGAGCTGCGCGCTGG 1872  
QY 6525 ACCTCGCGCGCCCAACCAAGTGACTCTTCACTTCGCCGCCCGGACCAATGGTGGCCCGGCC 6584  
Db 1873 ACCTCGCGCGCCCAACCAAGTGACTCTTCACTTCGCCGCCCGGACCAATGGTGGCCCGGCC 1932  
QY 6585 GTGGCTCCCCCAAGCGGAAGGGGGAAGAGTCTTCAGAGCCAAACAGACTGTGCT 6644  
Db 1933 GTGGCTCCCCCAAGCGGAAGGGGGAAGAGTCTTCAGAGCCAAACAGACTGTGCT 1992  
QY 6645 TGGGATGATGAGGACGGTATTGAACCTGTGTCCCAACCGGAGGCAATGACGAGCCAG 6704  
Db 1993 TGGGATGATGAGGACGGTATTGAACCTGTGTCCCAACCGGAGGCAATGACGAGCCAG 2052  
QY 6705 GGCATCTCCGAGTGTGTATACCCGCTGCTGTACCGGATGGGGAACAGACGAGCCCA 6764  
Db 2053 GGCATCTCCGAGTGTGTATACCCGCTGCTGTACCGGATGGGGAACAGACGAGCCCA 2112  
QY 6765 GCAGGATGGGCTCCAAAGTCTCCAGGCAACCAAGCCAGCCGCTTCTTCAAGAAC 6824  
Db 2113 GCAGGATGGGCTCCAAAGTCTCCAGGCAACCAAGCCAGCCGCTTCTTCAAGAAC 2172  
QY 6825 TGAACGAGAGCAATCCCGCAATGTCAGTCCAGAGCAAGAGATCAACAGAGCTGA 6884  
Db 2173 TGAACGAGAGCAATCCCGCAATGTCAGTCCAGAGCAAGAGATCAACAGAGCTGA 2232  
QY 6885 AACACCAACACGGAATGAGCTGTAATCAATATCAGCAGCCTGGAGCGAGATCTTCA 6944  
Db 2233 AACACCAACACGGAATGAGCTGTAATCAATATCAGCAGCCTGGAGCGAGATCTTCA 2292  
QY 6945 ATATGCCCCCATCAACCGGAACAGGCTTATGACTTATGAAAGCCAGGCGGTGACGAAAC 7004  
Db 2293 ATATGCCCCCATCAACCGGAACAGGCTTATGACTTATGAAAGCCAGGCGGTGACGAAAC 2352  
QY 7005 ATGCGAGACCAACATGGGGCTGAGGCAATTAAGAAAGGCACTCATGGTAAATATG 7064  
Db 2353 ATGCGAGACCAACATGGGGCTGAGGCAATTAAGAAAGGCACTCATG----- 2402  
QY 7065 ACAAGTGGAAAGTCCCGCGCTCAGGCGCAATGCTTTTAAACCTCTGAATGCAGTG 7124  
Db 2403 ----- 2402  
QY 7125 CCAAGCTGCGCGCTGATGCCATTAACCGCTGACGAGCGAGTGAACACACTCA 7184

Db 2403 ----- 2402  
QY 7185 CCGCCAGTGGCGGCGGAAGGCAAGTCTCTGGGAGACCCAGCAGCCGAAAGCCA 7244  
Db 2403 -----GTTGGCGGCGGAAGGCAAGTCTCTGGGAGACCCAGCAGCCGAAAGCCA 2454  
QY 7245 AGTCCCGGCGCGGAGCTGGATCTGGGGAACCGGCAACCTCTGTCTTCAAGTCACT 7304  
Db 2455 AGTCCCGGCGCGGAGCTGGATCTGGGGAACCGGCAACCTCTGTCTTCAAGTCACT 2514  
QY 7305 CGGAGGAGACTTGCACCGCGGAGACCGGCTCAACCAACCGGTGAGGAGACAGCCCT 7364  
Db 2515 CGGAGGAGACTTGCACCGCGGAGACCGGCTCAACCAACCGGTGAGGAGACAGCCCT 2574  
QY 7365 CGTCCGAGGTTCCAGGCAATTCCTTCAACCCCTGATCATGCGGCTGACAGCGGGTG 7424  
Db 2575 CGTCCGAGGTTCCAGGCAATTCCTTCAACCCCTGATCATGCGGCTGACAGCGGGTG 2634  
QY 7425 TCATGCTTCCCAACCCCAACCGGCTCTCCCGCGGAGCGGGCCCTCGCTGGCCCC 7484  
Db 2635 TCATGCTTCCCAACCCCAACCGGCTCTCCCGCGGAGCGGGCCCTCGCTGGCCCC 2694  
QY 7485 ACCAAGCTTGGAGCGAGAGCCCAAGCCACTGTCTGCTGCAAGTACAGACACTTCCG 7544  
Db 2695 ACCAAGCTTGGAGCGAGAGCCCAAGCCACTGTCTGCTGCAAGTACAGACACTTCCG 2754  
QY 7545 ACAAGGATGACTCAGAACAGGCGGGGGGGCGGCGTTEUBERSPERLUTGAGTCC 7604  
Db 2755 ACAAGGATGACTCAGAACAGGCGGGGGGGCGGCGTTEUBERSPERLUTGAGTCC 2806  
QY 7605 CAGCGAGCCACAGGAACGGCTTGCAGAGCGGGGCGGCTGCCAATCCCCCAACCAAG 7664  
Db 2807 CAGCGAGCCACAGGAACGGCTTGCAGAGCGGGGCGGCTGCCAATCCCCCAACCAAG 2866  
QY 7665 AAGAGCCCCCTGAGTCCGCTGCTCCATCTCTCTCCGTCAGAGCCGGCATCTT 7724  
Db 2867 AAGAGCCCCCTGAGTCCGCTGCTCCATCTCTCTCCGTCAGAGCCGGCATCTT 2926  
QY 7725 GCGT 7728  
Db 2927 GCGT 2930

Search completed: March 12, 2006, 12:21:16  
Job time : 5998 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2006, 09:17:38 ; Search time 1329 Seconds  
(without alignments)

11463.868 Million cell updates/sec

Title: US-09-522-753-4

Perfect score: 8564.8  
Sequence: 1 catgctcggtccacacagc.....caaaaaaaaaaaaaa 8571

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents NA:\*  
1: /cgn2\_6/prodata/1/ina/1 COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5 COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/H COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/PCTUS COMB.seq:\*  
7: /cgn2\_6/prodata/1/ina/PP COMB.seq:\*  
8: /cgn2\_6/prodata/1/ina/RE COMB.seq:\*  
9: /cgn2\_6/prodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8226.8	96.1	9053	3	US-09-976-594-306 Sequence 306, App
2	8230	7.4	7912	3	US-09-949-016-3804 Sequence 3804, App
3	630	7.4	7940	3	US-09-632-0338-1 Sequence 1, App1
4	629.6	7.4	7780	3	US-09-632-0338-2 Sequence 2, App1
5	311.2	3.6	312	3	US-09-513-999C-33415 Sequence 33415, A
6	142	1.7	1922	2	US-08-372-652-9 Sequence 9, App1
7	142	1.7	1922	6	PCT-US95-16311-9 Sequence 9, App1
8	107.8	1.3	3489	2	US-08-728-323A-1 Sequence 1, App1
9	107.8	1.3	3489	3	US-09-298-568-1 Sequence 1, App1
10	107.8	1.3	3489	3	US-09-410-399-1 Sequence 1, App1
11	107.8	1.3	3489	3	US-09-894-273-1 Sequence 1, App1
12	107.8	1.3	32207	2	US-08-770-379-20 Sequence 20, App1
13	107.8	1.3	32207	3	US-08-757-669A-20 Sequence 20, App1
14	107.8	1.3	32207	3	US-09-330-371A-20 Sequence 20, App1
15	104.6	1.2	187595	3	US-09-949-016-15546 Sequence 15546, A
16	96.2	1.1	1926	3	US-09-249-585A-2 Sequence 2, App1
17	96.2	1.1	1926	3	US-09-410-399-3 Sequence 3, App1
18	96.2	1.1	2580	3	US-09-050-863-2 Sequence 2, App1
19	96.2	1.1	2580	3	US-09-359-081-2 Sequence 2, App1
20	96.2	1.1	5452	2	US-09-130-114-1 Sequence 1, App1
21	96.2	1.1	8705	3	US-09-647-344A-14 Sequence 14, App1
22	96.2	1.1	9600	3	US-08-910-647-1 Sequence 1, App1
23	96.2	1.1	9600	3	US-09-620-925-1 Sequence 1, App1
24	96.2	1.1	10596	2	US-07-884-811-15 Sequence 15, App1

25	96.2	1.1	10596	2	US-07-885-971-15 Sequence 15, App1
26	96.2	1.1	10596	2	US-08-087-783A-15 Sequence 15, App1
27	96.2	1.1	10596	2	US-08-194-088B-15 Sequence 15, App1
28	96.2	1.1	10596	2	US-08-194-087-15 Sequence 15, App1
29	96.2	1.1	10596	6	PCT-US93-04648-15 Sequence 15, App1
30	96.2	1.1	16080	3	US-09-724-566A-48 Sequence 48, App1
31	96.2	1.1	16080	3	US-09-471-669A-48 Sequence 48, App1
32	94.8	1.1	127	3	US-09-680-420A-17 Sequence 17, App1
33	94.8	1.1	601	3	US-09-949-016-135614 Sequence 135614, A
34	93	1.1	94	3	US-09-513-999C-29269 Sequence 29269, A
35	92	1.1	1852	3	US-09-969-852-4 Sequence 4, App1
36	90	1.1	1926	3	US-09-249-585A-4 Sequence 4, App1
37	90	1.1	1931	2	US-09-130-114-2 Sequence 2, App1
38	89.8	1.0	7218	2	US-08-232-463-14 Sequence 14, App1
39	86.2	1.0	1995	2	US-08-425-069-3 Sequence 3, App1
40	86.2	1.0	1995	2	US-08-317-844B-3 Sequence 3, App1
41	85.4	1.0	16442	3	US-08-781-891-208 Sequence 208, App
42	85.4	1.0	16442	3	US-09-618-166-208 Sequence 208, App
43	82.8	1.0	767677	3	US-09-949-016-12147 Sequence 12147, A
44	82.8	1.0	767677	3	US-09-949-016-17361 Sequence 17361, A
45	82.4	1.0	2093	3	US-10-104-047-1666 Sequence 1666, Ap

#### ALIGNMENTS

RESULT 1  
US-09-976-594-306  
; Sequence 306, Application US/09976594  
; Patent No. 6673549  
; GENERAL INFORMATION:  
; APPLICANT: Furness, Michael  
; APPLICANT: Buchinder, Jenny  
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
; FILE REFERENCE: PA-0041 US  
; CURRENT APPLICATION NUMBER: US/09/976, 594  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: 60/240, 409  
; PRIOR FILING DATE: 2000-10-12  
; NUMBER OF SEQ ID NOS: 1143  
; SOFTWARE: PERL Program  
; SEQ ID NO 306  
; LENGTH: 9053  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Incyte ID No. 6673549 898877.6  
; NAME/KEY: unsure  
; LOCATION: 2006, 2012  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-976-594-306

Query Match 96.1%; Score 8226.8; DB 3; Length 9053;

Best Local Similarity 98.3%; Pred. No. 0;  
Matches 8451; Conservative 2; Mismatches 42; Indels 103; Gaps 10;

1 CATGTCGGGCTCCACACAGCTTGTGCAACAGCTGAGAGGCGCACTGAGCCCGCTAACCC 60  
|||||  
519 CATGTGGGATCCACACACAGCTGTGGCAAGAGTGAAGGCGCACTGAGCCCGCTAACCC 578  
|||||  
61 GCGCCACAGCTTCTTACCGAGTGCAGATCGCCGAGCGACACGAGCGTGGGCTCCT 120  
|||||  
579 GCGCCACAGCTTCTTACCGAGTGCAGATCGCCGAGCGACACGAGCGTGGGCTCCT 638  
|||||  
121 GAGATACGACGACCACTCCGCGACATATGCTCCACCTGCGCGGGCTCATATCCA 180  
|||||  
639 GAGATACGACGACCACTCCGCGACATATGCTCCACCTGCGCGGGCTCATATCCA 698  
|||||  
181 GCGCCACGCGGAGGCGCTTCTGCTGTGAGTTCCAGCCCGGAGATGACGCTTCCA 240  
|||||  
699 GCGCCACGCGGAGGCGCTTCTGCTGTGAGTTCCAGCCCGGAGATGACGCTTCCA 758  
|||||

QY 241 GAGACTCCAGCTGGGCGAAGTCCCACTCATACTCTCCCGAGCTGGGAACTCAGAGAT 300  
Db 759 GAGAGTCCACCTGGCGGCGAAGTCCCACTCATACTCTCCCGAGCTGGGAACTCAGAGAT 818  
QY 301 GAGAGTTCATTGAAACCAAGCCGCTTCGGGTAGAGCTGCTGCTGACCCCTCTGCTGAC 360  
Db 819 GAGAGTTCATTGAAACCAAGCCGCTTCGGGTAGAGCTGCTGCTGACCCCTCTGCTGAC 878  
QY 361 GTCAACCCCTGCTGGGCGAAGGCTGGGATCTGAAAGACTCAACCAAGACCGTAG 420  
Db 879 GTCAACCCCTGCTGGGCGAAGGCTGGGATCTGAAAGACTCAACCAAGACCGTAG 938  
QY 421 CTTGACGGGCGAAGCTGGAACCGGTGTCTTCCCGCCAGCCCCCGGCACTGACCTTGAGCT 480  
Db 939 CTTGACGGGCGAAGCTGGAACCGGTGTCTTCCCGCCAGCCCCCGGCACTGACCTTGAGCT 998  
QY 481 GAGAGTGGTCCGCGCAAGGCTGTCCAGAGAGAGCTGATCCAGAACATGGAACCGGTGGA 540  
Db 999 GAGAGTGGTCCGCGCAAGGCTGTCCAGAGAGAGCTGATCCAGAACATGGAACCGGTGGA 1058  
QY 541 CCGAGAGATCAACATGATGAGAGCAGACAGATCTTAAAGCTGAAAGAAAGCAACAGCT 600  
Db 1059 CCGAGAGATCAACATGATGAGAGCAGACAGATCTTAAAGCTGAAAGAAAGCAACAGCT 1118  
QY 601 GAGAGAGAGGCTGCGCAAGCCGCGGAGCTGAGAAAGCCGTGTCAACGCGCGCCATGGA 660  
Db 1119 GAGAGAGAGGCTGCGCAAGCCGCGGAGCTGAGAAAGCCGTGTCAACGCGCGCCATGGA 1178  
QY 661 GTCCAGAGACCGGCAAGCTGTGTCAGATCATCTTACAGACGAAACCGGAAAGAGCTGAAGC 720  
Db 1179 GTCCAGAGACCGGCAAGCTGTGTCAGATCATCTTACAGACGAAACCGGAAAGAGCTGAAGC 1238  
QY 721 TGCAATCGGATTTGGAAGGCTTGGGGCCCAAGGTGAGCTGCCGTGTACAAACAGCC 780  
Db 1239 TGCAATCGGATTTGGAAGGCTTGGGGCCCAAGGTGAGCTGCCGTGTACAAACAGCC 1298  
QY 781 CTCCGACACCCGGCGAGTATCATGAGAACATCAAAATTAACCAAGGATGCGGAAGAGCT 840  
Db 1299 CTCCGACACCCGGCGAGTATCATGAGAACATCAAAATTAACCAAGGATGCGGAAGAGCT 1358  
QY 841 AATCTTTGATCTTCAAGAGAGAGAAATCAAGCTTCGGAACAAATGGAAGCAGAAATTTCTGCCA 900  
Db 1359 AATCTTTGATCTTCAAGAGAGAGAAATCAAGCTTCGGAACAAATGGAAGCAGAAATTTCTGCCA 1418  
QY 901 GCGCTATGACCAAGCTCATGAGAGGCTTGGAAAAAAAGGTGAGCCGATCGAAAAACAACC 960  
Db 1419 GCGCTATGACCAAGCTCATGAGAGGCTTGGAAAAAAAGGTGAGCCGATCGAAAAACAACC 1478  
QY 961 GCGCGGCGGGGCGAAGAGAGCAAGGTGCGGAGTACTAGAAAAACAAGTTCCCTGAGAT 1020  
Db 1479 GCGCGGCGGGGCGAAGAGAGCAAGGTGCGGAGTACTAGAAAAACAAGTTCCCTGAGAT 1538  
QY 1021 CCGGAAGACGCGGAGCTGCAAGAGACGATGCAAGAGGCTGGGCGCAAGCGGGGCGAGTGG 1080  
Db 1539 CCGGAAGACGCGGAGCTGCAAGAGACGATGCAAGAGGCTGGGCGCAAGCGGGGCGAGTGG 1598  
QY 1081 GCTGTCCATGTGGGCGCGCGGAGGAGCAAGAGGTGTCAGAGATCATGAGTGGCTCTTC 1140  
Db 1599 GCTGTCCATGTGGGCGCGCGGAGGAGCAAGAGGTGTCAGAGATCATGAGTGGCTCTTC 1658  
QY 1141 AGAGCAGAGAAACTTGAAGAGCAAGTGGCGCAGCTGGGCTGGATCCGCGCCATGCTGTA 1200  
Db 1659 AGAGCAGAGAAACTTGAAGAGCAAGTGGCGCAGCTGGGCTGGATCCGCGCCATGCTGTA 1718  
QY 1201 CGAGCCTGACCAAGCGGCTCATAGTTCTAAACATGAAACGGGCTTTATGGCGGACCCCAT 1260  
Db 1719 CGAGCCTGACCAAGCGGCTCATAGTTCTAAACATGAAACGGGCTTTATGGCGGACCCCAT 1778  
QY 1261 GAAAGTGTACAAAGACCGCAGAGTATGAAACATGTGAGTGAAGAGAGAAAGAAACCTT 1320  
Db 1779 GAAAGTGTACAAAGACCGCAGAGTATGAAACATGTGAGTGAAGAGAGAAAGAAACCTT 1838  
QY 1321 CCGGAGAAAGTTTCATGACAGCATCCAGAAACTTTGGCCTGATCGATCATCTTCTGAGAG 1380

Db 1839 CCGGAGAAAGTTTCATGACAGCATCCCAAGACTTTGGCCTGATCGATCATCTTCTGAGAG 1898  
QY 1381 GAAAGCAGTGGCTGAGTGGGCTCTCTATTAATCACTGAGCTAAGAGAAATGAGAACTATTA 1440  
Db 1899 GAAAGCAGTGGCTGAGTGGGCTCTCTATTAATCACTGAGCTAAGAGAAATGAGAACTATTA 1958  
QY 1441 GAGCCTGTGAGACGAGACTATTCGCGCGCGGCGCAAGAGCCAGCAGCAACA- --CAGCA 1497  
Db 1959 GAGCCTGTGAGACGAGACTATTCGCGCGCGGCGCAAGAGCCAGCAGCAACACAGCA 2018  
QY 1498 GAGAGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1557  
Db 2019 GAGAGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2078  
QY 1558 GAAAGATGAG 1617  
Db 2079 GAAAGATGAG 2138  
QY 1618 GAAACGACAAAGAGAGACTCTTCAAGAGAGAGACAGACGACCTCAGGAGAGAGCAACGA 1677  
Db 2139 GAAACGACAAAGAGAGACTCTTCAAGAGAGAGACAGACGACCTCAGGAGAGAGCAACGA 2198  
QY 1678 GAGAGAGAGGCTGTGGCTCTCAAGAGCCGCAAACTGACCAACAGCCAGGAGAGCGCA 1737  
Db 2199 GAGAGAGAGGCTGTGGCTCTCAAGAGCCGCAAACTGACCAACAGCCAGGAGAGCGCA 2258  
QY 1738 AGGCGCATCAACCGGCTCAATGGCTAATGAGGCGCAACGAGAGAGGCGCATCACCCCA 1797  
Db 2259 AGGCGCATCAACCGGCTCAATGGCTAATGAGGCGCAACGAGAGAGGCGCATCACCCCA 2318  
QY 1798 GCAAGCCGCGAGCTGAGCTCCATGAGAGCTGATGAGAGTTCTCGCTGACAGAAAGAA 1857  
Db 2319 GCAAGCCGCGAGCTGAGCTGAGCTCCATGAGAGCTGATGAGAGTTCTCGCTGACAGAAAGAA 2378  
QY 1858 AATGAGAAACAGCCAAAGAAAGTCTCTTGAACACGCGCGCACTGTGGGCTGATGGCCCG 1917  
Db 2379 AATGAGAAACAGCCAAAGAAAGTCTCTTGAACACGCGCGCACTGTGGGCTGATGGCCCG 2438  
QY 1918 GATGGTGGGCTCCAAAGCTGTGTGCAAGTGTAAAGAACTTCACTCACTCAACAAAGAG 1977  
Db 2439 GATGGTGGGCTCCAAAGCTGTGTGCAAGTGTAAAGAACTTCACTCACTCAACAAAGAG 2498  
QY 1978 GCAGAACTTCATGAGATCTTTCAGCAGCAGCAACAGCTGAGAAATGAGAGAGAGAGAGC 2037  
Db 2499 GCAGAACTTCATGAGATCTTTCAGCAGCAGCAACAGCTGAGAAATGAGAGAGAGAGAGC 2558  
QY 2038 GCGAG 2096  
Db 2559 GCGAG 2618  
QY 2097 TGGTGAAGATGAGAGAGATGAGAGGCTGGGCGTGAAGGCGAAATGAGAGAGAGATGAGTGG 2156  
Db 2619 TGGTGAAGATGAGAGAGATGAGAGGCTGGGCGTGAAGGCGAAATGAGAGAGAGATGAGTGG 2678  
QY 2157 AGGAGGCTGAAGCCTTACATGCTCTCTGGAAATGAGGTGCCAGAGGGAGATGCACTGCGC 2216  
Db 2679 AGGAGGCTGA----- 2688  
QY 2217 CAGGCACTGTCAACACAGCTTCAGACACCGAGAGCATCCCTCTCTCACTAGAGGCG 2276  
Db 2689 AGGCACTGTCAACACAGCTTCAGACACCGAGAGCATCCCTCTCTCACTAGAGGCG 2747  
QY 2277 CCAAGGACACAGGGAGAGATGGGCGCCAGGCGCCAGCAGCCTGCGGCGCCAG- GGGGCA 2335  
Db 2748 CCAAGGACACAGGGAGAGATGGGCGCCAGGCGCCAGCAGCCTTGGGCGCGAGCGGGGCA 2807  
QY 2336 CCCCCAGGCCACCAACCCCAACAGAGAGCATTCGCGGCGCCCATTTAGGCCACCCCG 2395  
Db 2808 CCCCCAGGCCACCAACCCCAACAGAGAGCATTCGCGGCGCCCATTTAGGCCACCCCG 2867  
QY 2396 GCTCTGAAGCCACCGAGGCCCTTACGCCCCCAACAGCAACCCCATCGCCCTTGCACCT 2455

Db	2868	GCCTGTGAAGCACCAGAGCCCTTACGCCCCCAACAGACCCCCATCGCCTCTGCACCT	2927
OY	2456	CCTCTGTGTGTCCCAAGAGAGAGAGAGAGACCGAGCAGCGCCCCAGTGTAG	2515
Db	2928	CCTCTGTGTGTCCCAAGAGAGAGAGAGAGACCGAGCAGCGCCCCAGTGTAG	2987
OY	2516	GAGGGGGAGAGACAGAACCCCCCGGCTTGAAGACTGTGGCAATGTGACACAGGGAAAGCC	2575
Db	2988	GAGGGGGAGAGACAGAAACCCCCCGGCTTGAAGACTGTGGCAATGTGACACAGGGAAAGCC	3047
OY	2576	GAGAGAGCCCGTCAAGAGAGAGTGACACGAGAGAAACCGAGAGGGGGCCGCAAGGCAAG	2635
Db	3048	GAGAGAGCCCGTCAAGAGAGAGTGACACGAGAGAAACCGAGAGGGGGCCGCAAGGCAAG	3107
OY	2636	GACCGGAGAGCCGCTGTAGGCCACCGGCCAGGGGGCGCTCAAGGCAAGAGAGAGAGGC	2695
Db	3108	GACCGGAGAGCCGCTGTAGGCCACCGGCCAGGGGGCGCTCAAGGCAAGAGAGAGAGGC	3167
OY	2686	GAGAGCCGAGAGGCCACCACTGACCAAGCTCGGGCGCCCCAGAGACAGGACCTCCAGT	2755
Db	3168	GAGAGCCGAGAGGCCACCACTGACCAAGCTCGGGCGCCCCAGAGACAGGACCTCCAGT	3227
OY	2756	GCTACCTGACATGACAGAGAGGTGGATGAGGCCGAGGGGGCGCAAGAAACGGCTCTGT	2815
Db	3228	GCTACCTGACATGACAGAGAGGTGGATGAGGCCGAGGGGGCGCAAGAAACGGCTCTGT	3287
OY	2816	TCCCAAGAGGCCACGACTCTCAACCCCGACTGAGCACCCCGGGGCAATGTCTTCAACCCAG	2875
Db	3288	TCCCAAGAGGCCACGACTCTCAACCCCGACTGAGCACCCCGGGGCAATGTCTTCAACCCAG	3347
OY	2876	AAGCCACTGACCTTGAGACGCTGAGACGAGCGAGCGCTGCAATCCCCCCTACAGATC	2935
Db	3348	AAGCCACTGACCTTGAGACGCTGAGACGAGCGAGCGCTGCAATCCCCCCTACAGATC	3404
OY	2936	ACCAAGTCCATGAGCCCCCCCCGGAGAGAGCAGACTCCCAACAGCCAGCTCCCCCAG-C	2994
Db	3408	ACCAAGTCCATGAGCCCCCCCCGGAGAGAGCAGACTCCCAACAGCCAGCTCCCCCAGAC	3464
OY	2995	CCCAACCGGCACCGCAAAACCTGACACCGAGAGAGAGCGCCCCCTGAGAGACTGTGACAG	3054
Db	3465	CCCAACCGGCACCGCAAAACCTGACACCGAGAGAGAGCGCCCCCTGAGAGACTGTGACAG	3524
OY	3055	CCCCCGGGGCAAGAGACGAGACCCCGGACCCCGCCGCAAGAGAG-----	3100
Db	3525	CCCCCGGGGCAAGAGACGAGACCCCGGACCCCGCCGCAAGAGAGAGAGACTGT	3584
OY	3101	-----GCCTTCGATCGCCAGAGGCCCAAGATGCTGTGGGAGACCCCTTGTCTGTAC	3150
Db	3585	GTTCTTCCAGACCTTTCGAGCGCCAGAGGCCCAAGATGCTGTGGGAGACCCCTTGTCTGTAC	3644
OY	3151	TTCCGGACCTGTGCTTCCCGCGAGCCCCCGAGAGGTGATCAAGGCTTCCCGCATGTGCCCC	3210
Db	3645	TTCCGGACCTGTGCTTCCCGCGAGCCCCCGAGAGGTGATCAAGGCTTCCCGCATGTGCCCC	3704
OY	3211	GGAACCCCTCAACCTTCTCTTCAAGCTGTCACTGTGTACCCCATGACCTGCGGCTTCAATGA	3270
Db	3705	GGAACCCCTCAACCTTCTCTTCAAGCTGTCACTGTGTACCCCATGACCTGCGGCTTCAATGA	3764
OY	3271	CACGTGCCCGGCGGCTCTGCGCGCGCCACCAACATCTCCAAACCGGCTCCCTCATCTCC	3824
Db	3765	CACGTGCCCGGCGGCTCTGCGCGCGCCACCAACATCTCCAAACCGGCTCCCTCATCTCC	3884
OY	3331	CTCTGACCAAGACCCCAAGCGTCTCTGAGAGGCAATATGTGGCAATCTCCAGGAATGTCT	3390
Db	3825	CTCTGACCAAGACCCCAAGCGTCTCTGAGAGGCAATATGTGGCAATCTCTCCAGGAATGTCT	3884
OY	3391	GATCAGACTCAGCTCCCGTACTGACAGCAATGCAAGGCCCCGGTGGGACCTGTCAACAT	3450
Db	3885	GATCAGACTCAGCTCCCGTACTGACAGCAATGCAAGGCCCCGGTGGGACCTGTCAACAT	3944
OY	3451	GGGGGTGCGCTCGCCCATGTGACACCCCAAAACCTGTGACCTTTACGCGAGTGAAGCAGGA	3510
Db	3945	GGGGGTGCGCTCGCCCATGTGACACCCCAAAACCTGTGACCTTTACGCGAGTGAAGCAGGA	4004

OY	3511	CCAGCTGTCCTCCACGGGGCCAGGCTGGGCAACCGAGACCTGGGGGTGCCACAGCCCA	3570
Db	4005	GCAGCTGTCCTCCACGGGGCCAGGCTGGGCAACCGAGACCTGGGGGTGCCACAGCCCA	4064
OY	3571	GGAGGCGTCCGGTCTGAGAGGAGCAGCTTGGGCTTCAGTTCCTGGGCGGAAAGATCAACA	3630
Db	4065	GGAGGCGTCCGGTCTGAGAGGAGCAGCTTGGGCTTCAGTTCCTGGGCGGAAAGATCAACA	4124
OY	3631	AGGCATTTCCCAACACACGGGGTGCCTTCGGACAGCGCATATCACTATCCGCGCTCCATCAC	3690
Db	4125	AGGCATTTCCCAACACACGGGGTGCCTTCGGACAGCGCATATCACTATCCGCGCTCCATCAC	4184
OY	3691	CCACGGCACGCGACGCTGACGTCCTGTACAAAGGACCATCAACAGGATCATCGGCGAGA	3750
Db	4185	CCACGGCACGCGACGCTGACGTCCTGTACAAAGGACCATCAACAGGATCATCGGCGAGA	4244
OY	3751	CAGCCCGAGTCGCTTGGACCCGCGCCGGAGAGACAGCTTCCCAAGGAGCCCATCTTA	3810
Db	4245	CAGCCCGAGTCGCTTGGACCCGCGCCGGAGAGACAGCTTCCCAAGGAGCCCATCTTA	4304
OY	3811	CGAAGGCAAGAAAGGGCCACGTCCTTGTCCATATAGGTTGGCATGTCTGTGACCCAGTCTC	3870
Db	4305	CGAAGGCAAGAAAGGGCCACGTCCTTGTCCATATAGGTTGGCATGTCTGTGACCCAGTCTC	4364
OY	3871	CAAGGAGACGGCAGAACAGCTCAGACCCGCCCATGAGACGCGCCGCCCAAGGCGCAC	3930
Db	4365	CAAGGAGACGGCAGAACAGCTCAGACCCGCCCATGAGACGCGCGCCGCCCAAGGCGCAC	4424
OY	3931	CTATGACATGATGAGGGCCGCGTGGGAGAGGCATTTCTTCAGCCACGATCGAAGTCT	3990
Db	4425	CTATGACATGATGAGGGCCGCGTGGGAGAGGCATTTCTTCAGCCACGATCGAAGTCT	4484
OY	3991	CATGGGCGGTCGCAATCCGCGCGGAGACGACACAGCCGCCACACCTCAAAAGAGACCA	4050
Db	4485	CATGGGCGGTCGCAATCCGCGCGGAGAGACGACACAGCCGCCACACCTCAAAAGAGACCA	4544
OY	4051	CATCCGCGGGTCCATCACAAAGGATCCCTCGTCTTACGTGAGGCAACAGGAGACTA	4110
Db	4545	CATCCGCGGGTCCATCACAAAGGATCCCTCGTCTTACGTGAGGCAACAGGAGACTA	4604
OY	4111	CCTCGCTGGGAGGCGCAAGCTCTTAACCGGAGGACACGCTCCGCCCCACCGGCTTC	4170
Db	4605	CCTCGCTGGGAGGCGCAAGCTCTTAACCGGAGGAGGACGCTCCGCCCCACCGGCTTC	4664
OY	4171	ACGGGACCTTGACCGAGGCTCTAACAGCGACGCGCTGGGCGCCCTTGAACTGAAAGCGGCG	4230
Db	4665	ACGGGACCTTGACCGAGGCTCTAACAGCGACGCGCTGGGCGCCCTTGAACTGAAAGCGGCG	4724
OY	4231	CCATGAGGGCTGTGGTGCCACGCTGTAAGAGAGGCGGCGGCTCCATTCATGAGATCCCGCG	4290
Db	4725	CCATGAGGGCTGTGGTGCCACGCTGTAAGAGAGGCGGCGGCTCCATTCATGAGATCCCGCG	4784
OY	4291	CGAGGAGCTGGGGACACGCGCCGAACGTGCGCCCTGGCGCGGCGGCGCTCAAGAGGGCTC	4350
Db	4785	CGAGGAGCTGGGGACACGCGCCGAACGTGCGCCCTGGCGCGGCGGCGCTCAAGAGGGCTC	4844
OY	4351	CATCACCGAGGGCACCCCGCTCAAGTACGACACCGGCGGCTCACCACTGGCTCAAAAA	4410
Db	4845	CATCACCGAGGGCACCCCGCTCAAGTACGACACCGGCGGCTCACCACTGGCTCAAAAA	4904
OY	4411	GCACGACGTACGCTTCCCTCATATGGGACGCCCGGAGCGTTTCCACCCGTGCAACCGCT	4470
Db	4905	GCACGACGTACGCTTCCCTCATATGGGACGCCCGGAGCGTTTCCACCCGTGCAACCGCT	4964
OY	4471	GGATGTGATGCGCACGCGCCGGGCACTGGAAGCTGCTGTCTGAGAGAGGCTGAAAGAG	4530
Db	4965	GGATGTGATGCGCACGCGCCGGGCACTGGAAGCTGCTGTCTGAGAGAGGCTGAAAGAG	5024
OY	4531	CCGGCCAGGGACCGCCAGACGTCGGGGGGCTCAATTGCGCGGGCGCCCGGTCAATTGT	4590
Db	5025	CCGGCCAGGGACCGCCAGACGTCGGGGGGCTCAATTGCGCGGGCGCCCGGTCAATTGT	5084

QY	4591	CCCTGAGCTGGGGTAAAGCCGGGCAAGAGCCCTGACCTA	TAAGACAACAGGGGACCTT	4650
Db	5085	GCCTGAGCTGGGCAAGCCGGGCAAGAGCCCTTA	CTATAGGACAACAGGGGACCTTT	5144
QY	4651	TGCCGGCACCTCCACGAGGGTTGGCCCTGATG	ACATGGGGAGGCCACGCCGCGCTTGCA	4710
Db	5145	TGCCGGCACCTCCACGAGGGTTGGCCGTTG	ACACAGCGGGAGGCCACGCCGCGCTTGCA	5204
QY	4711	GGAGGGCAGCCTTTTCGTCCAGCAAGGCATCC	CAAGACTGACGTGACGCTCG	4770
Db	5205	GGAGGGCAGCCTTTTCGTCCAGCAAGGCATCC	AGAGCCGAAAGCTGACGTGACGCTCG	5264
QY	4771	TGAGATGCGCAAGTCCCGGCAACAGACCGT	AGCCAGACCAACCCCATCTGGCC	4830
Db	5265	TGAGATGCGCAAGTCCCGGCAACAGACCGT	AGCCAGACCAACCCCATCTGGCC	5324
QY	4831	CTATAGACCACTGCTGGGGCGGTGAGTGG	GGTGGACCTGATCGACACATCCCTT	4890
Db	5325	CTATAGACCACTGCTGGGGCGGTGAGTGG	GGTGGACCTGATCGACACATCCCTT	5384
QY	4891	GGCCTTGACCCCACTTCATACCCCGGCGAT	CTCTGAGAGCAAGCGCTGCTACTA	4950
Db	5385	GGCCTTGACCCCACTTCATACCCCGGCGAT	CTCTGAGAGCAAGCGCTGCTACTA	5444
QY	4951	CTGCCCCCGAACCCTGGCCCCCAACCCCA	CTTACCCGACCTGTACCAACCTTACCTAT	5010
Db	5445	CTGCCCCCGAACCCTGGCCCCCAACCCCA	CTTACCCGACCTGTACCAACCTTACCTAT	5504
QY	5011	CCGGGGCTACCCCGACACAGCGGGCTTGA	GAACCGGACACATCATATGACTACAT	5070
Db	5505	CCGGGGCTACCCCGACACAGCGGGCTTGA	GAACCGGACACATCATATGACTACAT	5564
QY	5071	CACCTCGCAGATGACCAACAACCGGCCA	CCGCAATGGCCCAAGCGAGCTGATGCT	5130
Db	5565	CACCTCGCAGATGACCAACAAGCGGCCA	CCGCAATGGCCCAAGCGAGCTGATGCT	5624
QY	5131	GAGGGGCTCTTCGCCCCCGACAGTCTCG	CTGGACTCAACTACGTGGGGTCCCCCAGG	5190
Db	5625	GAGGGGCTCTTCGCCCCCGACAGTCTCG	CTGGACTCAACTACGTGGGGTCCCCCAGG	5684
QY	5191	CATCATCGACTGTGCCAAGTGCACACCTG	CTGTGCTGGGCCCCCGACACCAAGGAC	5250
Db	5685	CATCATCGACTGTGCCAAGTGCACACCTG	CTGTGCTGGGCCCCCGACACCAAGGAC	5744
QY	5251	CCAGCCACCGCACATGACCGGCTTGCTTA	CCCTCCACCGCGGCCCAAGCCCTTCAAGAG	5310
Db	5745	CCAGCCACCGCACATGACCGGCTTGCTTA	CCCTCCACCGCGGCCCAAGCCCTTCAAGAG	5804
QY	5311	CCGCCACAGCAGCTTCCCATCTTCCCGAG	AGGTTCACAACATTTGACAAAACCAACAC	5370
Db	5805	CCGCCACAGCAGCTTCCCATCTTCCCGAG	AGGTTCACAACATTTGACAAAACCAACAC	5864
QY	5371	CAGCTCCCTCGCCGAGCGGGAGCGAGAC	CCGGAGTCCAGAGGGGACCGGGATCCGGGAGCG	5430
Db	5865	CAGCTCCCTCGCCGAGCGGGAGCGAGAC	CCGGAGTCCAGAGGGGACCGGGATCCGGGAGCG	5924
QY	5431	GGAAAAGTCATCTCTCACTCAACACGAG	CGGTGAGACCGCACCCATCTTGAGACCTGG	5490
Db	5925	GGAAAAGTCATCTCTCACTCAACACGAG	CGGTGAGACCGCACCCATCTTGAGACCTGG	5984
QY	5491	TACAGAGCAGACAGCGGCAAGCGGCGA	CGGCGGGGTGGGGCAGACAGACGCG	5550
Db	5985	TACAGAGCAGACAGCGGCAAGCGGCGA	CGGCGGGGTGGGGCAGACAGACGCG	6044
QY	5551	CCCGGCGTCCCATCTCCATATGCCCCA	ACAGCATTCGCCCTGGGACCCAGAGATGC	5610
Db	6045	CCCGGCGTCCCATCTCCATATGCCCCA	ACAGCATTCGCCCTGGGACCCAGAGATGC	6104
QY	5611	CCTCGACGAGACCAAGTGTGCTTCAACA	ACAGGCAATGAAGGATATCAACAACGCTGT	5670
Db	6105	CCTCGACGAGACCAAGTGTGCTTCAACA	ACAGGCAATGAAGGATATCAACAACGCTGT	6164
QY	5671	GGAGCCCAAGCCACAGTCTTG-----	AGTCCACTTCACCTTCTCAACCGGT	5721

Db	6165	GGAGCCCAAGCAGCGCCACGGTCTCGAGGTGGGCCAGGTCCACCTCCACCTCTCACCCGT	6224
Qy	5722	TCGCGCAGCTGCACATATCCCACTGTGCACCTCACTGCGGCCACCTCGATGG	5781
Db	6225	TCGCGCGGCTGCACATATCCCACTGTGCACCTCACTGCGGCCACCTCGATGG	6284
Qy	5782	GGTTACCTTAACCTTCATTTGAGGCCGCTTTGGCTGGCCCAAGAGGCCCGCGGGTCGCGG	5841
Db	6285	GGTTACCTTAACCTTCATTTGAGGCCGCTTTGGCTGGCCCAAGAGGCCCGCGGGTCGCGG	6344
Qy	5842	GCCAGACGCGGCCCGAGACAGACCCGGCCATATGCTTCTTGCCAAAGCCCCCAGCGCGCTC	5901
Db	6345	GCCAGACGCGGCCCGAGACAGACCCGGCCATATGCTTCTTGCCAAAGCCCCCAGCGCGCTC	6404
Qy	5902	CGGGCTGAGAGCCCGCTCTCTCCCCAGACAAAGGCTCGAGAGCCCCCGCCCTAGTGCCTC	5961
Db	6405	CGGGCTGAGAGCCCGCTCTCTCCCCAGACAAAGGCTCGAGAGCCCCCGCCCTAGTGCCTC	6464
Qy	5962	TGTCTCTGGGCACGCGCACATATGCGCGCACCTCTGCGAAGAACTTGCACTTACCAACGC	6021
Db	6465	TGTCTCTGGGCACGCGCACATATGCGCGCACCTCTGCGAAGAACTTGCACTTACCAACGC	6524
Qy	6022	CAGCCCGGACCCCGCGCGCACCTGTGCTGCGACTCGGACCCGACCGGAAAAAGACTCA	6081
Db	6525	CAGCCCGGACCCCGCGCGCACCTGTGCTGCGACTCGGACCCGACCGGAAAAAGACTCA	6584
Qy	6082	AAGTAAACCTTTTCCATCCAGAACTGAACTTCGTTCTGCGGTTACCAAGCGACAG	6141
Db	6585	AAGTAAACCTTTTCCATCCAGAACTGAACTTCGTTCTGCGGTTACCAAGCGACAG	6644
Qy	6142	CTTACAGCCCCGAGAGGGGTGAGGCCCTGTACGCTGTAGACTCAACCACTTGAACCAACA	6201
Db	6645	CTTACAGCCCCGAGAGGGGTGAGGCCCTGTAGACTCAACCACTTGAACCAACA	6704
Qy	6202	CAAGGGGCTCCCCAAGACCTGGAAGGCTCGACAAGAGGCCACTGGAGGGGAGCTGCG	6261
Db	6705	CAAGGGGCTCCCCAAGACCTGGAAGGCTCGACAAGAGGCCACTGGAGGGGAGCTGCG	6764
Qy	6262	GCCCAACAGCAGAGGCCCGCTGTAAGCTTTGGCGGGAGGCCGCCACCTTCCACACTGCG	6321
Db	6765	GCCCAACAGCAGAGGCCCGCTGTAAGCTTTGGCGGGAGGCCGCCACCTTCCACACTGCG	6824
Qy	6322	GCCGCTGCTTGAGAGCCAGGCCCTTGTGCAGGCCGCTGTCCAGACCGCCCCAGGGGTCAA	6381
Db	6825	GCCGCTGCTTGAGAGCCAGGCCCTTGTGCAGGCCGCTGTCCAGACCGCCCCAGGGGTCAA	6884
Qy	6382	AGGTCAACAACGGGTGGTTCACTCCGGGCCAGCACTATGAGAGGTCACTCAACAGACTA	6441
Db	6885	AGGTCAACAACGGGTGGTTCACTCCGGGCCAGCACTATGAGAGGTCACTCAACAGACTA	6944
Qy	6442	CACCCGGCACCAACCAAGAGACTCAGCGGCAACCCCTGCGCCGCCCTCTTAATCTTTCCC	6501
Db	6945	CACCCGGCACCAACCAAGAGACTCAGCGGCAACCCCTGCGCCGCCCTCTTAATCTTTCCC	7004
Qy	6502	TGGGGCAGGTGCCCGCTCTGTAACCTTCGCGCGCCACCCAGTGAACCTTATCTCCGCC	6561
Db	7005	TGGGGCAGGTGCCCGCTCTGTAACCTTCGCGCGCCACCCAGTGAACCTTATCTCCGCC	7064
Qy	6562	CCCGGAACAATGTATCCCCCGGCTCCCTCCCAACAGGAAGGGGGCAAGAGTCTCC	6621
Db	7065	CCCGGAACAATGTATCCCCCGGCTCCCTCCCAACAGGAAGGGGGCAAGAGTCTCC	7124
Qy	6622	AGAGCCAAACAAAGACTCGGCTTTTGGGTGTGTGTGAGAGCGTATTAACTTGTCCTCC	6681
Db	7125	AGAGCCAAACAAAGACTCGGCTTTTGGGTGTGTGTGAGAGCGTATTAACTTGTCCTCC	7184
Qy	6682	ACCAGAGGGCATGACGAGGCCAAGGCACTCCCGAGTGTGTGTATCCCGTGTCTGACCG	6741
Db	7185	ACCAGAGGGCATGACGAGGCCAAGGCACTCCCGAGTGTGTGTATCCCGTGTCTGACCG	7244
Qy	6742	GGATGGGAAACAGCGGAGCCCAAGAGATGGGTCTCCAATCTTCAAGCAACACCACTCA	6801



7245 GATGGGGAACAGACGAGCCAGAGATGGCTCCAGCTCTCCAGGCAACAGCCCA 7304  
6802 GCCGACGACCTTTTGAAGCAAGCTGACCCGAGAGCAATCCGCCATGGTCAAGTCCAGAA 6861  
7305 GCCGACGACCTTTTGAAGCAAGCTGACCCGAGAGCAATCCGCCATGGTCAAGTCCAGAA 7364  
6862 GCAAGAGATCAACAGAGAGTGAACCCCAACCCGAGATGAGCTGAAATACAAATACAG 6921  
7365 GCAAGAGATCAACAGAGAGTGAACCCCAACCCGAGATGAGCTGAAATACAAATACAG 7424  
6922 CCAAGCTGGGAGCGAGATCTTCAATATGCCCCGATCACTGAGAAAGGCTTTATGACCTA 6981  
7425 CCAAGCTGGGAGCGAGATCTTCAATATGCCCCGATCACTGAGAAAGGCTTTATGACCTA 7484  
6982 TAGAAGCCAGGCGGTGAGGAAATGCAAGCAATGAGGCTGAGAGGCTGAAATTTAG 7041  
7485 TAGAAGCCAGGCGGTGAGGAAATGCAAGCAATGAGGCTGAGAGGCTGAAATTTAG 7544  
7042 AAGGCACTCATGGGTAAATATGACCAAGTGGGAAAGTCCCGCGCTCAGCGCCCAATGC 7101  
7545 AAGGCACTCATGGGTAAATATGACCAAGTGGGAAAGTCCCGCGCTCAGCGCCCAATGC 7604  
7102 TTTTAAACCTCTGATATGCAAGTCCAGCTGCGCTGCTATATGCCAATTAACCGCTGCTGA 7161  
7605 TTTTAAACCTCTGATATGCAAGTCCAGCTGCGCTGCTATATGCCAATTAACCGCTGCTGA 7664  
7162 CGGACGAGATGACCACTCACTCGCAGGTGCGCGGGAAGGCGCAAGGCTCTCG 7221  
7665 CGGACGAGATGACCACTCACTCGCAGGTGCGCGGGAAGGCGCAAGGCTCTCG 7724  
7222 CAGACCCAGACGCGGAAAGCAAGTCCCGCGCGCGGCTGAGCACTTGGGAGACCGGAC 7281  
7725 CAGACCCAGACGCGGAAAGCAAGTCCCGCGCGCGGCTGAGCACTTGGGAGACCGGAC 7784  
7282 ACCCTGTGTCTCTCAATGCACTCGAGGAGAGTCAAGCAACCGCGCTCAACCA 7341  
7785 ACCCTGTGTCTCTCAATGCACTCGAGGAGAGTCAAGCAACCGCGCTCAACCA 7844  
7342 CGCGGTGAGGAGAGAGGCGCTCGTCCGAGGTTCAAGCCATTTCCCTCAACACCCCT 7401  
7845 CGCGGTGAGGAGAGAGGCGCTCGTCCGAGGTTCAAGCCATTTCCCTCAACACCCCT 7904  
7402 GATCATGCGGCTGCAAGCGGAGTGTCAAGCTTCCCAACCCGAGCTCCCGCGAG 7461  
7905 GATCATGCGGCTGCAAGCGGAGTGTCAAGCTTCCCAACCCGAGCTCCCGCGAG 7964  
7462 CAGCGGCGCTCTGCTGCGGCGCCCAACCGCTGAGGAGAGCCCAAGCTGCTTG 7521  
7965 CAGCGGCGCTCTGCTGCGGCGCCCAACCGCTGAGGAGAGCCCAAGCTGCTTG 8024  
7522 CTCGCGATCGAGACCTCTCCGACAGCGAGTCACTCAAGACGCGGCGGCGGCGGCGG 7581  
8025 CTCGCGATCGAGACCTCTCCGACAGCGAGTCACTCAAGACGCGGCGGCGGCGGCGG 8084  
7582 CGGTUERAERLUJTCAGGTCCCAAGCAAGCAAGAAAGGCGCTGCAAGAGCGGAGCG 7641  
8085 -----GCGGTGTCAGGTCCCAAGCAAGCAAGAAAGGCGCTGCAAGAGCGGAGCGG 8135  
7642 GCTGCGACTCCCGCAACCAAGAGAGCGCTGAGTCCGCTGCGCTCATTCATCT 7701  
8136 GCTGCGACTCCCGCAACCAAGAGAGCGCTGAGTCCGCTGCGCTCATTCATCT 8195  
7702 GTCCGTCAGAGCGCGGATCTTGTCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 7761  
8196 GTCCGTCAGAGCGCGGATCTTGTCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAG 8255  
7762 CTGGCCCTGTGCAAGCTTAACTCAAGGAGATTTTAACTGGTGTCCGGAAGGAGGAGAA 7821  
8256 CTGGCCCTGTGCAAGCTTAACTCAAGGAGATTTTAACTGGTGTCCGGAAGGAGGAGAA 8315  
7822 GGGGCGGAGAGGAGGAGCGGAGCGGTGTGAGGAGCAACAGGCGGCGAGGCGGCG 7881  
8316 GGGGCGGAGAGGAGGAGCGGAGCGGTGTGAGGAGCAACAGGCGGCGAGGCGGCG 8375

7882 AGGACCCAAAGCAGATGACCAAGCACTTCAAGCACTGCTCCCGGATGATTTG 7941  
8376 AGGACCCAAAGCAGATGACCAAGCACTTCAAGCACTGCTCCCGGATGATTTG 8435  
7942 GAACCAAGCTTAACTGAGCTGCAAGCGCGCGCTTCCCTCCCTCCATCCCGCT 8001  
8436 GAACCAAGCTTAACTGAGCTGCAAGCGCGCGCTTCCCTCCCTCCATCCCGCT 8495  
8002 TAGCGCTGAGACAGATGAGAGCGAGCGCTGTCAGCGCCGAGTGGCTGCTCCGCTC 8061  
8496 TAGCGCTGAGACAGATGAGAGCGAGCGCTGTCAGCGCCGAGTGGCTGCTCCGCTC 8555  
8062 CCACAGCTGCGCCAGCAGCAAGAGATTTGAGAAACCAAGTCAGGCGAGTGGCGAGCA 8121  
8556 CCACAGCTGCGCCAGCAGCAAGAGATTTGAGAAACCAAGTCAGGCGAGTGGCGAGCA 8615  
8122 AAGGCGCAGGTGCGGCTGCGGAGGAAAGGATGCTCGAGACTGGAATGTTTTTTCAC 8181  
8616 AAGGCGCAGGTGCGGCTGCGGAGGAAAGGATGCTCGAGACTGGAATGTTTTTTCAC 8675  
8182 ACATCGTTCGCGAGCGGTGGGAGAGGAGAGGATGTAATGATGTTGTTTACAG 8241  
8676 ACATCGTTCGCGAGCGGTGGGAGAGGAGAGGATGTAATGATGTTGTTTACAG 8735  
8242 GTATATTTTGAATACCTTCAATGAATTAATTCAGATGTTTTTTCAGAAAGGACTTACC 8301  
8736 GTATATTTTGAATACCTTCAATGAATTAATTCAGATGTTTTTTCAGAAAGGACTTACC 8795  
8302 CAGTATTAATGCTGCTGCTGCTTGTATCTCTGCTTCAAGAGCGGTGAGAGGCT 8361  
8796 CAGTATTAATGCTGCTGCTGCTTGTATCTCTGCTTCAAGAGCGGTGAGAGGCT 8855  
8362 GACAGTCGAGTACCCCATCACTCGAGAGCAAGAGGCGGAGAGCTGCTGCTGCTGCTG 8420  
8856 GACAGTCGAGTACCCCATCACTCGAGAGCAAGAGGCGGAGAGCTGCTGCTGCTGCTG 8915  
8421 CGCTGTGCTCTCCCTCCCTCCCTTCCCTTGGGAGAAATGATGATGATGATGATGATG 8480  
8916 CGCTGTGCTCTCCCTCCCTCCCTTCCCTTGGGAGAAATGATGATGATGATGATGATG 8975  
8481 CGCATTTGCGAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 8540  
8976 CGCATTTGCGAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 9035  
8541 GAATTTACTCCAAAAA 8558  
9036 GAATTTACTCCAAAAA 9053

RESULT 2  
US-09-949-016-3804  
; Sequence 3804, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: PasteSeq for Windows Version 4.0  
; SEQ ID NO 3804  
; LENGTH: 7912  
; TYPE: DNA  
; ORGANISM: Human

US-09-949-016-3804

Query Match 7.4% Score 630; DB 3; Length 7912;  
 Best Local Similarity 65.1%; Pred. No. 4.2e-108; Indels 48; Gaps 4;  
 Matches 1015; Conservative 0; Mismatches 495; Indels 48; Gaps 4;

491 CCGCCACGGCTGTCAGAGAGAGCTGATCCAGAACATGACCCGGCTGAGCCGAGAGATC 550  
 757 CTTCAAAACCTCTCAAGAGAGAGTTAATACAGAGATGATCGTGTAGATCGAGAAATT 816  
 551 ACCATGTGTAGACACAGATCTCTAAAGCTGAAAGAAAGACAGCAACAGCTGGAGAGAG 610  
 817 GCAGAAAGTAGAAGACAGACAGATCTTAAACCTGAAAAAGAAACAACAGCTTGAAGAG 876  
 611 GCTGCCAAGCCGCGGAGCTGAGAGAGCCGCTGTACCGCGCCATGAGTCGAGAGC 670  
 877 GCAGCTAAACCTCTGAGAGCTGAGAGAGCCGCTGTCTCTCTGTGTGAGCAGAAAC 936  
 671 CGCAGCCTGTGTGAGATCATCTACGAGAGAACCGAGAGAGAGCTGAAAGCTGACATCGG 730  
 937 CGCAGTATTTGTCAAATATTATTAATGATGAGATCGGAAAAAGAGAGAGCTCATTA 996  
 731 ATTCTGGAAGGCTGTGGGCCCCAGAGTGAAGCTGCGCTGTACACAGCCCTCCGACAC 790  
 997 ATTTTGAAGGCTTGTGGCCCAAGATTGAACTGCACTGTATTAACAGCCATCAGATTAC 1056  
 791 CGGAGATATCATGAGAACATCAAAATTAACAGAGGAGTCGAGAGAGCTAATCTTGTAC 850  
 1057 AAGGTGTACATGAGAACATCAAGAACAGAGTGTAGAGAGAGAGAGAGAGAGAGAG 1116  
 851 TTCAAGAGAGAGAGATCAGGCTCGAGAACATGAGAGAGAGAGAGAGAGAGAGAGAG 910  
 1117 TTTAAAGAGAGAGATCATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1176  
 911 CAGCTCATGAGAGAGCTTGTGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 970  
 1177 CAGCTCATGAGAGAGCTTGTGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1236  
 971 GCCAAG 1030  
 1237 GCTTAAG 1296  
 1031 CGCAGCTGTGAG 1090  
 1297 AG 1353  
 1091 TCGGCGGCGCGGAG 1150  
 1354 ACCATTTGTAG 1413  
 1151 AACCTGAG 1210  
 1414 AATATGTAG 1473  
 1211 CAGCAGCGCATCAAGTTCATCAACATGAACGGGCTTAATGCGGAGCCAGATGAAGTGTAC 1270  
 1474 CAAG 1533  
 1271 AAGAGCGGAG 1330  
 1534 AAGAGTAG 1593  
 1331 TTATGAG 1390  
 1594 TTTATCGAG 1653  
 1391 GCTGAGCGCTCTCTAATTAATCACTGAATGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1450  
 1654 CTTGATGTGTTTGTATTAATTAATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1713  
 1451 AG 1510  
 1714 AG 1765

1511 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1570  
 1766 CGCAG 1825  
 1571 GAG 1630  
 1826 AAG 1878  
 1631 GACCTCTCAAG 1690  
 1879 AATACAG 1938  
 1691 GTGCTCTCAAG 1750  
 1939 GCACACACCCCGGAG 1998  
 1751 CGCTCATGCTTAATGA-----GGCCACAGCAG 1780  
 1999 AGGTCTCATGAG 2058  
 1781 GAGGCTATCACCCCGAG 2118  
 2059 CCCCACCACTCTGCAACCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2178  
 1841 CGCTGAG 1900  
 2119 CGATGAG 2178  
 1901 TGTGCGCATGCGCCGAG 1960  
 2179 TGGGAG 2238  
 1961 TTCACTCAAG 2018  
 2239 TTTAATTAAG 2296

## RESULT 3

US-09-632-033B-1

Sequence 1, Application US/09632033B

Patent No. 6949624

GENERAL INFORMATION:

APPLICANT: Liu, Johnson M.

APPLICANT: Wang, Jianxiang

TITLE OF INVENTION: CLONING OF THE HUMAN NUCLEAR RECEPTOR

FILE REFERENCE: NIH172.001A

CURRENT APPLICATION NUMBER: US/09/632,033B

CURRENT FILING DATE: 2000-08-09

PRIOR APPLICATION NUMBER: 60/146,977

PRIOR FILING DATE: 1999-08-03

NUMBER OF SEQ ID NOS: 9

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 7940

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURES:

OTHER INFORMATION: cDNA encoding human nuclear receptor co-Repressor

OTHER INFORMATION: (Hun-Cor)

US-09-632-033B-1

Query Match 7.4% Score 630; DB 3; Length 7940;  
 Best Local Similarity 65.1%; Pred. No. 4.2e-108; Indels 48; Gaps 4;  
 Matches 1015; Conservative 0; Mismatches 495; Indels 48; Gaps 4;

491 CCGCCACGGCTGTCCAGAGAGAGCTGATCCAGAACATGACCCGGCTGAGCCGAGAGATC 550  
 757 CTTCAAAACCTCTCAAGAGAGAGTTAATACAGAGATGATCGTGTAGATCGAGAAATT 816  
 551 ACCATGTGTAGACACAGATCTTAAAGCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 610

Db 817 GCAAAAGTAAACAGACATCTTAAACTGAAAAAGAAACAACAAGCTTTGAAGAAG 876  
Qy 611 GGTGCAAGCCGCGCAGCTGAGAACCCCTGTCAACCCGCGCCATTCGATGGAAGC 670  
Db 877 GCAAGTAAACCTCTGAGGCTGAGAACCCGCTGTCCCTCTCTGTGAGCAAGAAC 936  
Qy 671 CCGAGCTGTGAGATCTCTACGACGAAACCGGAGAAAGGCTGACATCGG 730  
Db 937 CGCAGTATGTCAAAATTTATGATGAGAAATCGAAAAAGCAAGAAAGCTCATAA 996  
Qy 731 ATTCTGAAAGGCTGTGGGCCCCAGGTGAGCTGCGCTGTACAACAGCCCTCGACCC 790  
Db 997 ATTTTGAAGGTCTTGGCCCAAAAGTTGAACCTGCTATTAACAGCCATCAGATACC 1056  
Qy 791 CCGAGTATCATGAAACATCAAAATTAACAGGCAATGCGGAGAAAGCTAACTTGTAC 850  
Db 1057 AAGGTACCATGAAACATCAAGAACCAAGTATGAGAAAAAACTATTTATTT 1116  
Qy 851 TTCAAGAGAGAAATCAAGCTCGGAAACAATGGAACAGAACTTGCACGCTATGAC 910  
Db 1117 TTTAAAGAAAGAAATCATGCAAGAAACAAGGGAACAAAAATCTGCCAGCCTTATGAT 1176  
Qy 911 CAGCTCATGAGGCTTTGAAAAAAAGTGAAGCGATCGAAAAACAACCCGCGCGCG 970  
Db 1177 CAGCTCATGAGGCAATGGAGAAAAAAGTGAACAGAAATGAATAATCTCTCGAGGAAA 1236  
Qy 971 GCCAAGAAAGCAAGGTGCGCGATCTACGAAACAGTTCCTGTGATCCGCAAGCAG 1030  
Db 1237 GCTAAAGAAAGCAAAACAAGGAATCTGTGAAACAGTTTCAAAATTCGAAAAACA 1296  
Qy 1031 CCGAGCTCGAGAGGCAATGACAGACAGGATGGGCGAGCGGAGGCTGTCTCATG 1090  
Db 1297 AAGAAACAGCAAGAAAGATTTGAG---CGAGTTGGGCAAGGGAGACTGTCTTACGC 1353  
Qy 1091 TCGGCGCGCGCAGCGACGACAGAGTGTCAAGATCATGATGCTTCAAGACAGAG 1150  
Db 1354 ACCATGTGATGAGAGTGAAGATGATTTCTGAATATTATGATGGGCTCTCTGAGCAG 1413  
Qy 1151 AACCTGGAAGACAGATGCGCCAGCTGGCGGTATCCCGCCATGCTGTAACAGCCTGAC 1210  
Db 1414 AATTAATGAACAAATGCGGCGCTCTGTGATTTCCACTATGATGTTGATGAGAA 1473  
Qy 1211 CAGACCGCATCAAGTTTCAATCAATGAACGAGCTTATGGCCGACCCCATGAAGTGTAC 1270  
Db 1474 CAAGAAGAGTCAAGTTTCAATCAATGAATGGCTTATGAGAGACCTTATGAAGTGTAT 1533  
Qy 1271 AAAGACCGCAGGTCAATGAACATGTGAGTGAAGCAGAGAGAGACCTTCCGAGAG 1330  
Db 1534 AAAGATAGGCAATTTATGATGTTTGAATGACCATGAAGAGAGATCTTTAAGACAA 1593  
Qy 1331 TTCAATGACATCCCAAGAACTTTGGCTGTATGATCATTTCTGAGAGAGAAAGTGT 1390  
Db 1594 TTTATTCAGACATCAAAAACTTTGACCTAATTCATCACTTGAAGAGAGAGTGT 1653  
Qy 1391 GCTGAGTGTCTCTTAATTAACCTGATTAAGAAATGAAGAACTATTAAGAGCTGTGT 1450  
Db 1654 CTTGATGTGTTTGTATTAATTAATTAACAGAAATTAAGAAATTAAGAGCTGTGT 1713  
Qy 1451 AGACGAGCTATTCGCGCGCGCGCAAGACCAAGCAGCAACACAGCAGCAGCAG 1510  
Db 1714 AAGAGAAATTAATGAGAAAGCAGAGGCAAAACCAACA-----AATGCTCGACCT 1765  
Qy 1511 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1570  
Db 1766 CCGAAGAAAGAAAGTGAAGAAAGAAAGAGATTAACAGCAAAAAACAAGAAAAAG 1825  
Qy 1571 GAGAGAGAAAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1630  
Db 1826 AAGAGAGAA-----GAAAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1878  
Qy 1631 GACCTTCTCAAG 1690  
Db 1879 AATACCAAG 1938

Qy 1691 GTGCTCTCCAAAGGCGCGCAAAAGTGCACCAAGCCAGAGGAGAGAGAGAGAGAG 1750  
Db 1939 GCCACACCCCGGGGGGAGAAAGACTGCCAACAGTCAAGGGCCCGCTAAGGCCGATACC 1998  
Qy 1751 CGCTCAATGCTTAATGA-----GGCCAAACAGCAG 1780  
Db 1999 AGGTCCATGACAAACGAAGGTGAGCTGCGCAAGTGTGACAGCCGACGGCTTATGAAGAG 2058  
Qy 1781 GAGGCCATCACCCCCAGCAGAGAGCGCGAGCTGCTTCAATGAGCTGAATGAGTTCT 1840  
Db 2059 CCCCCACACCTCTGACCGCAGCAGCAACCCATTTTCAACAGAGCCTGTGAGAGACTCT 2118  
Qy 1841 CGCTGACAG 1900  
Db 2119 CGATGACAG 2178  
Qy 1901 TGGTGGCCATCGCCCGAGTGGTGGCTCAAGACTGTGTGAGTGAATCTTAC 1960  
Db 2179 TGGGCAAGCAATGTCTAAATGGTGGAGAGAAAGTGAAGCTCAATGTAAATCTTAT 2238  
Qy 1961 TTCACTAAG 2018  
Db 2239 TTTAACTATAAAGGCGACACATCTTGAACAACCTCTTACAGCAGCATAAACAGAAA 2296

RESULT 4  
US-09-632-033B-2  
; Sequence 2, Application US/09632033B  
; Patent No. 6949624  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Johnson M.  
; TITLE OF INVENTION: CLONING OF THE HUMAN NUCLEAR RECEPTOR  
; FILE REFERENCE: NIH172.001A  
; CURRENT APPLICATION NUMBER: US/09/632,033B  
; CURRENT FILING DATE: 2000-08-09  
; PRIOR APPLICATION NUMBER: 60/146,977  
; PRIOR FILING DATE: 1999-08-03  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 780  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: CDNA encoding murine nuclear receptor co-receptor  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(780)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-632-033B-2

Query Match 7.4%; Score 629.6; DB 3; Length 780;  
Best Local Similarity 64.4%; Pred. No. 4.9e-108;  
Matches 1029; Conservative 0; Mismatches 524; Indels 45; Gaps 4;

Qy 446 TCTCCCCCAGGCCCCCGGACACTGACCTGAGCTGAGACTGTGCGGACGCTGTCC 505  
Db 588 TCTCCCTCTCTGGGAGGCGATGAGATGATGAGATGCTTCAAAAGCTGTA 647  
Qy 506 AAGAGAGACTGATCCAGAACTGACCGCGTGAAGCCGAGAGATCAACATGTGAGCAG 565  
Db 648 AAGAGAGACTGATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 707  
Qy 566 CAGATCTTAAGCTGAAG 625  
Db 708 CAGATCTTAAGCTGAAG 767  
Qy 626 GAGCTGAG 685

```
Db 768 GAGCTGAGAGGCTGTGTCCCTCCCTCCCTGAGAGAGAGACCGAAGTATTTGCCAA 827
Qy 686 ATCATCTACGACGAGAACCGGAGAGAGGCTGAGAGTGCACATCGGATTTCTGGAGAGCCCTG 745
Db 828 ATCATTTTATGACGAGAAATCGGAAAAAGAGAGAAAGCTCAAAAATTTTGAAGTCTT 887
Qy 746 GGGCCCCAGGTGAGCTGCGCTGTACAAACAGCCCTCCGACACCCGGCAGTATCATGAG 805
Db 888 GGGCCCAAAAGTTGAACCTGCGCTTACCAACGCGCTGAGATACCAAGGTGATACACAG 947
Qy 806 AACATCAAAATTAACAGCCGATGCGGAAAGAGCTAATTTTACTTTCATTGAAGAGAGAT 865
Db 948 AACATCAAGACCAACAGGTGATGAGAAAAAATCTCATTTTATTTTAAAAAGAAAT 1007
Qy 866 CACGCTCGGAAACAATGAGACAGAGTCTGCCAGCGCTATACAGCTCATGAGAGCC 925
Db 1008 CATGCAAGAAAAACAAGGGAACAAAAATCTGCCAACGTTATGATCAGTCTATGAGAGA 1067
Qy 926 TTGAAAAAAGGTGAGCGCATCGAAAAACAACCGCGCGCGGCGCCAGAGAGACAG 985
Db 1068 TGGAGAAAAAAGTGGACAGATAGAAAAATATCTCGGAGGAAAGCTTAAAGAACAA 1127
Qy 986 GTGGCGGAGTACTAGAAAAAGAGTTCCCTGAGATCCGCAAGCAGCGAGCTGAGAG 1045
Db 1128 ACAGAGGAATACTATGAAAAAGCAGTTTCCAGAAATTCGAAAAACAAGAGACAGAGA 1187
Qy 1046 CGCATGACAGAGCAGGATGCGGCGCAGCGGCGAGTGGGCTGTCCATGTGCGCCGCCAGC 1105
Db 1188 AGATTTTCAG---CGAGTTGTGTAGAGGGAGGAGTGGTCTTTTCAGCCACCATTTGAGAGT 1244
Qy 1106 GAGCAGAGGTGTCAAGATCATCGATGCGCTCTCAAGACAGAGAACTGTGAGAGCAG 1165
Db 1245 GAGCATGAGATTTCTGAATTTATGTGTCTTTCTGAAACAGAGAAATATGAGAGCA 1304
Qy 1166 ATGGCCGAGCTGGCGGTATCCCGCCCATGCTGTACAGCGTCAAGAGAGGAGCATCAAG 1225
Db 1305 ATGCTCTAGCTTTCGTGATTTCCACCTATGATGTTTGAAGCAAAAGAGGCTCAA 1364
Qy 1226 TTCACTCAACATGAAAGGCTTATGGCCGACCCCATGAAAGGTATCAAAAGCCGACAGTC 1285
Db 1365 TTCACTCAATATGAAATGGGCTGATGAGATCCAAATGAAGTTTATTAAGACAGACAGTTT 1424
Qy 1286 ATGAACATGTGAGTGAAGAGAGAGAGACCTTCCGGAGAGGTTCAAGAGCATCC 1345
Db 1425 ATGAATGTTTGAACATGATAAAAGAGATCTTTAAGACAGAGTTTATCAGATCCA 1484
Qy 1346 AAGAACTTTGGCTGATCGCATCATCTCTGAGAGAGAAAGACGTGGAGTGGCTCTC 1405
Db 1485 AAAAACTTTGACCTAATTTGCATCTTATTTGAAAGAGAAAGTTCCTGATTTGTGTTTA 1544
Qy 1406 TATTACTACTGACTAAGAAAGATGAAGAACTATTAAGAGCTGTGTGAAGCGAGTATCGG 1465
Db 1545 TATTACTATTTAACCAAGAAAAATGAAATTTAAGGCCCTGTGTGAAGAAATTTATGA 1604
Qy 1466 CGCCCGCGAGAGCCAGCAGCAACAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1525
Db 1605 AAACCGCAGAGGCGAATATCAGCA-----GATTCGCGTCCCTCAACAAGAG 1650
Qy 1526 CAGCAGCCCATGCCCCCGCAGCAGCCAGAGAGAGAAAGATGAGAGAGAGAGAAAGAGAG 1585
Db 1651 AAAAAAGTAAAGAAAAAGAAAGAGATTAAGCAGAAAAACAG-AGAAAAAGAGAAAGAA 1709
Qy 1586 GCGGAGAGAGAGAGAGAGAGCCGAGAGGTGAGAAAGCAAGAGAGAGAGCTCTCAAGAG 1645
Db 1710 AAGAAAGATGATGAAGAAAAATATGATTAAGAGAGACTTTAAGAGAAACAACAGAGAAAG 1769
Qy 1646 AAGACAGACGACCTCAGGGAGAGCAACGAGAGAGAGAGCTGTGGCTCTCAAGAGC 1705
Db 1770 GACAGAGCGGAGCCAGCAGCAGAGAACTGAAAGAAAGAGACAGGTCACTCAAGGGGG 1829
Qy 1706 CGCAAAATTCGCAACGCGGAGAGAGAGCAAAAGCCGATCATCCGCTCAATGGCTAAT 1765
Db 1830 CGAAAGACTGTAAACAGCCAGCGCGGGAGAGGCGGGGTCAACAGGTGATGACAACT 1889
```

```
Qy 1766 GAG-----GCACAAGAGAGAGGCCATCAACCCCCAG 1798
Db 1890 GAGGTGACAGCTGCCAATGTGTCTGTCTGCACCTCAGTGAAGAGCCCCGCCACCCCTGCGG 1949
Qy 1799 CAGAGCCCGAGAGTGGCTCCATGAGAGCTGAATGAGATTTCTGCTGGAACAGAAAGAA 1858
Db 1950 CCACCACAGAGCCCATTTTTCAGAAACCTTTGAGCTTGAAGATGAGACAGAAAGAA 2009
Qy 1859 ATGAAACAGCCAAAGAAAGTCTCCTGGAACAGCGCCGCAACTGGTGGCCATCGCCGG 1918
Db 2010 ATGAAAGTTGCTTAAAAAGGCGCTGGTAGAATGTGTCTTAAGTGGCAGCCATTTCTAAA 2069
Qy 1919 ATGGTGGCTCCAAAGCTGTGTGCGAGTGAAGAACTTTCTAATTCACTCAAGAGAGG 1978
Db 2070 ATGGTGGGAACTTAAAAAGTGAAGCCAGTGCAGAAAACTTATTTTAACTATTAAGACGG 2129
Qy 1979 CAGAACTTCGATGAGATCTTTCAGAGACAGCAAGCTGAA 2016
Db 2130 CATATCTTGAACAACCTTTTGCAGCAACATTAACAGAA 2167
```

```
RESULT 5
US-09-513-999C-33415
; Sequence 33415, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent, pm
; SEQ ID NO 33415
; LENGTH: 312
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 86
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 87
; OTHER INFORMATION: m=a or c
US-09-513-999C-33415
```

Query Match 3 6%; Score 311.2; DB 3; Length 312;  
Best Local Similarity 99.4%; Pred. No. 6.3e-49;  
Matches 310; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1560 AAGATGAGAGAGAGAGAGAGAGAGGCGGAGAGAGAGAGAGAGAGCGGAGGTGAGAG 1619
Db 1 AAGATGAGAGAGAGAGAGAGAGAGAGGCGGAGAGAGAGAGAGAGAGCGGAGGTGAGAG 60
Qy 1620 ACGACAGAGAGAGCTCTCTCAAGAGAGAGACGACACTCAGGGAGAGCAACGAGCG 1679
Db 61 ACGACAGAGAGAGCTCTCTCAAGAGAGAGAGACGACACTCAGGGAGAGCAACGAGCG 120
Qy 1680 AGAAGAGGCTGTGGCTCTCAAGAGCGCGCAAACTGCGCAACGCGTCAAGAGAGAGCGCAAG 1739
Db 121 AGAAGAGGCTGTGGCTCTCAAGAGCGCGCAAACTGCGCAACGCGTCAAGAGAGAGCGCAAG 180
Qy 1740 GCGGATCAACCGCTCAATGGCTAATGAGGCCAACAGGAGAGAGGCTCATCCCCCAGC 1799
Db 181 GCGGATCAACCGCTCAATGGCTAATGAGGCCAACAGGAGAGAGGCTCATCCCCCAGC 240
```



TELEPHONE: 617/542-5070  
 TELEFAX: 617/542-8906  
 TELEEX: 200154  
 INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1922 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA  
 PCT-US95-16311-9

Query Match 1.7%; Score 142; DB 6; Length 1922;  
 Best Local Similarity 52.5%; Pred. No. 4.4e-17;  
 Matches 435; Conservative 0; Mismatches 375; Indels 18; Gaps 5;

6730 GCTGCTGATACCGGAGTGGGAAACAGACGAGCCGACGAGGAGTGGCTCCAACTCTCCAG 6789  
 1039 GCTCTTGTCACAGAGGAGGAGTGGACCTGCTGAGCAAGAGAGTGAATTCGATCAACAGG 1098  
 6790 CAACACGAGCCGAGCCGAGCCTCTTTCAGCAAGCTGACCGAGCAACTCCGCAATGGT 6849  
 1099 AAGTATAGCTACTTGCCTTCAATCTTTCACCAAGCT---TGAAGCAGATCAACCATGGT 1155  
 6850 CAAGTCCAGAAAGCAGAGATCAACAGAGCTGAAACCCCAACCGGATGAGCTTGA 6909  
 1156 TAAATCAAGAAACGAGAAATTTTCTGAAGTTGAATCTTCTGGTGGAGGTGACTTGA 1215  
 6910 ATACAAATTCAGCCGCTGGGACGAGATCTTCAATATGCCCCGATCAACCGGAACGG 6969  
 1216 TATGGCAGCTGCTCAAGAGGAGGAGTCTTCAATCTGCAAGAGTGAATCAACATCAGG 1275  
 6970 CTTATGACTATAGAGCCAGGCGGTGCGAGACATGACCAACCATGAGGCTGGA 7029  
 1276 TGAGTGAAGCTCAAGAGCATCTTTTGCTGATCCCGCAG---TAACTTGGTCTAGA 1332  
 7030 GGCATTAATTGAAAGGCACTCATGAGTAAAT---ATGACAGTGGGAAAGATCCCGCC 7086  
 1333 AGACATCATCAGAAAGGCTCTCATGGGAAGTTTGAATGATTAAGTTGAAGATCATGTGT 1392  
 7087 GCTCAGGCGCAATGCTTTTAACCCCTGAATGCCAGTGCACCTGCCCCGCTGCTATGCC 7146  
 1393 TGTCTATGTCCTCAATCTGTGGGATTAATGCTGTGTGTGTCACACCTCAGTGTACGAG 1452  
 7147 CATTAACCGCTGTGACGAGCGAGTGAACACACACTCTGCGCAGGTGGCGGGGGA 7206  
 1453 CAGCGAGGCAAGGAGATGAAAGGGAGCATCACTCATGAGAGATGTAACCAAA 1512  
 7207 GAGCAAGGCTCTTGGCAGACCCAGACGCGGAAAGCCAAATCCCGGCCCCGAGCTGGC 7266  
 1513 GCTGATCAACAATCAAAACAGAGGAAGCTTAAATCTCTATTCTGGGCAAAAGCTATTT 1572  
 7267 ATCTGGGAGACCGGCAACCTCTGTCTCTCTCAAGTGACTCGGAGGAGACTGAAACCGCG 7326  
 1573 AGGAATCAAAAGGCTTCTTCTGTCTCTCTGTGATTCAGAAAGGTGATTAACAAGGGA 1632  
 7327 GACGCGCTCAACAACCGGTGTGGAGGAGCAGGCGCTGTCGAGAGGTTCCAGCCATT 7386  
 1633 GACAC-----CAGATGGGATGGAGAGATGGGCGCTCTTCAACAGGTTCTACTCAATT 1686  
 7387 CCCCTTAACAACCCCTGATCATAGCGGCTGACGCGGGGTGTGATGCTTCCCAACCCCAAC 7446  
 1687 CCTTCAACAACCTCTGACCATATCGGATGCTCAGAGTACACCACTTACCAATGCGATG 1746  
 7447 GGGCTTCCCGGGGAGC---GGGCGCTGCGCTGGCCCCCAACAGCTGGGAAGAGA 7503  
 1747 GCGCCCATCTGCATACCTCAAGAGCTTCACATCAACGAACCCGATCTGGGAGAGGGA 1806  
 7504 GCCCAAGCAGCTGCTGCTGCTGAGTGAAGAGACACTCTCCGACAGGA 7551  
 1807 GCTGCGCCGCTCTCTCTCAGCGCAGTATGAGACACTGCTGTGATGTA 1854

# RESULT 8 US-08-728-323A-1

; Sequence 1, Application US/08728323A  
 ; Patent No. 5948676

; GENERAL INFORMATION:  
 ; APPLICANT: Chang, Yuan

; APPLICANT: Bohenzky, Roy A.

; APPLICANT: Russo, James J.

; APPLICANT: Edelman, Isidore S.

; APPLICANT: Moore, Patrick S.

; TITLE OF INVENTION: Immediate Early Protein From Kaposi's

; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:  
 ADDRESS: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/728,323A

FILING DATE:  
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
 NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-278-0400

TELEFAX: 212-391-0525

INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:

LENGTH: 3489 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:  
 NAME/KEY: CDS

LOCATION: 1..3489

US-08-728-323A-1

Query Match 1.3%; Score 107.8; DB 2; Length 3489;  
 Best Local Similarity 43.8%; Pred. No. 1.3e-10;

Matches 631; Conservative 0; Mismatches 792; Indels 18; Gaps 3;

219 AGCCGGGAATGAACGATGCCAGAGCTTCACCTGCGGCGAGATCCACTATACCTGC 278  
 1352 AGCCACAGCAGAGGAGCAGACGAGGAGGCCCTTCAGAGAGCCACAACAGCAGAGC 1411  
 279 CCGAGCTGGGGAAGTCAGAGATGAGTTCATTGAAGAAGCGCCCTGCGCTAGAGCTGC 338  
 1412 CACAGCAGAGAGGAGCAGACGAGAGGCCCTTCAGAGAGCCACAACAGCAGAGCAGC 1471  
 339 TGCCTGACCCCTGCTGGAACCTGTCACCTGCTGCGGCGAGCGGCGAGCTGCG 398  
 1472 AGCAGCAGAGAGCCCTTGAGAGGCCACAACAGCAGAGCCACAACAGCAGAGCAGC 1531  
 399 AAGACTTACCAAGAGCCGTAGCTGACGCGGCGAGCTGGAACCGGTGTCTCCCCCAGCC 458  
 1532 AGCAGAGCCACAGCA-----GCAGAGCCACAGCAGCAGAGGCCACAGCAGCAGAGC 1585  
 459 CCCGCACTGAGCCCTGAGCTGAGCTGTGCGGCGCAGCGCTGTCCAAGAGAGCTGA 518  
 1586 CACAGCAGAGAGCCACAGCAGCAGAGCCACAGCAGCAGAGGAGCAGCAGAGC 1645





Db 1886 AGCAGCAGCAGATGACGACGACGAGATGACGACGACGATGACGACGACGATG 1945  
QY 819 ACCAGCGATGCGAAGAACTAATCTTCTTCAAGAGAGAAATCAAGCTCGGAAC 878  
Db 1946 AGCAGCAGCAGATGACGACGACGAGATGACGACGAGATGACGACGAGATG 2005  
QY 879 AATGGAAGGAAAGTTCTCCGACCGCTATGACCAAGTTCATGAGCCCTTGGAAAAAG 938  
Db 2006 AGCAGCAGCAGATGACGACGACGAGATGACGACGACGAGATGACGACGAGATG 2065  
QY 939 TGGAGCGATCGAAAAACAACCCGCGCGGCGCCAAAGAGCAAGAGTGGCGAGTACT 998  
Db 2066 AGCAGCAGCAGATGACGACGACGAGATGACGACGAGATGACGACGAGATGACGACG 2125  
QY 999 ACGAAAGCAGTTCCCTGAGATCCGCAAGCAGCGCAGCTGACGAGCCGATGACAGCA 1058  
Db 2126 AGGATGAGCAG-----GAGCAGCAGAGATGACGACGACGAGATGACGACGAG 2179  
QY 1059 GGGTGGGCGAGCGGGGCACTGGGCTGTCCATGTCGCGCCCGCAGCCGACGACGAGTGT 1118  
Db 2180 ATGACGACGACGACGACGAGATGACGACGACGACGACGATGACGACGACGAGATGAGC 2239  
QY 1119 CAGAGATCATGATGAGCTCTGACGACGACGACGACGACGACGACGACGACGACGCTGG 1178  
Db 2240 AGCAGCAGCAGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2299  
QY 1179 CCGTATCCCGCCCATGCTGTGACGACGCTGACGACGACGACGACGACGACGACGACG 1238  
Db 2300 AGGATGATGAGAGCAGCAGCAGAGTTAGAGATCAGAGCAGAGGATTAGAGAGCAGG 2359  
QY 1239 ACGGCTTATGCGCCGACCCCATGAGGTTCACAAAGACCGCCAGGTCATGAACATGTGGA 1298  
Db 2360 AGCAGAGTTTACAGAGCAGCAGCAGAGTTTACAGAGCAGCAGCAGAGTTTACAGAGC 2419  
QY 1299 GTGAGCAGGAAAGAGACCTTCCGGGAGAGATTCAATGACGATCCCAAGACTTTG--- 1355  
Db 2420 AGGAGCAGAGATTAGAGAGCAGCAGCAGAGATTAGAGAGCAGAGCAGAGTTAGAGG 2479  
QY 1356 ---GCCCTGATCATATTCCTGAGAGAGAACAGTGGCTGATGTCCTCTATTACT 1412  
Db 2480 AGCAGAGCAGAGATTAGAGAGCAGCAGCAGAGATTAGAGAGCAGAGGTTAGAGAGC 2539  
QY 1413 ACCTGATTAAGAAATGGAATCTATTAGAGCCTGTGTGAGACGAGCTATCGGCCCGCG 1472  
Db 2540 AAGACAGAGGTTGAGAGAGCAAGCAGAGCAGAGCAGAGAAATTAGAGAGGTGG 2599  
QY 1473 GCAAGAGCAGAGCAACAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1532  
Db 2600 AGGACCAAGAGCAGAGCAGAGAGCAGAGAGAGAGAGATTAGAGAGGTGAAAGAC 2659  
QY 1533 CCATGCCCCGACGACCCAGAGAGAGAAAGATGAGAGAGAGAGAGAGAGAGAGAG 1592  
Db 2660 AGGAAGAGAGAGATTAGAGAGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGGTGG 2719  
QY 1593 AGGAG 1652  
Db 2720 AAGACAG 2779  
QY 1653 A 1653  
Db 2780 A 2780

RESULT 10  
US-09-410-399-1  
; Sequence 1, Application US/09410399  
; Patent No. 6482387  
; GENERAL INFORMATION:  
; APPLICANT: Robertson, Erle S.  
; APPLICANT: Corbett, Murray A.  
; TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA  
; TITLE OF INVENTION: to Genomic Host DNA  
; FILE REFERENCE: UM-03778

; CURRENT APPLICATION NUMBER: US/09/410,399  
; CURRENT FILING DATE: 1999-10-01  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3489  
; TYPE: DNA  
; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-410-399-1

Query Match 1.3%; Score 107.8; DB 3; Length 3489;  
Best Local Similarity 43.8%; Pred. No. 1.3e-10;  
Matches 631; Conservative 0; Mismatches 792; Indels 18; Gaps 3;

QY 219 AGCCCGGAAATGAACGATCCAGAGCTTCACTTGGCCAGAGTCCCATCTAATCTTC 278  
Db 1352 AGCCACAG 1411  
QY 279 CCGAGCTGGGAAAGTCAAGATGAGATTGATTAAGAAAGCAAGCGCCCTGGGCTAGAGCTGC 338  
Db 1412 CACAGCAGCAGAGAGCAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1471  
QY 339 TGGCTGACCCCTGCTGAGACCGTACCCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 398  
Db 1472 AGCAGCAGAGAGCTCTTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1531  
QY 399 AAGACTTCAACAG 458  
Db 1532 AGCAG 1585  
QY 459 CCGCCGACATGAGACCTTGAAGCTGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 518  
Db 1586 CACAGCAGCAG 1645  
QY 519 TCCAGAAATGAG 578  
Db 1646 CACAGCAG 1705  
QY 579 TGAAGAAAG 638  
Db 1706 CACAGCAG 1765  
QY 639 CCGTGTACCGCCGCTCATCGAGTGAAGAGCAGCAGCAGCTGTGTGAGATCATCTACAGG 698  
Db 1766 CACAGCAG 1825  
QY 699 AGAAGCGGAAAG 758  
Db 1826 AGCAGCAGAGAGATGAG 1885  
QY 759 AGCTGCGCTGTACCAACAG 818  
Db 1886 AGCAGCAGAGAGATGAG 1945  
QY 819 ACCAGGCGATGCGGAAAGAGTATCTTGTACTTCAAGAGAGAGAGATCAAGCTCGGAAC 878  
Db 1946 AGCAGCAGAGAGATGAG 2005  
QY 879 AATGGAAGGAAAGTTCTCCGACCGCTATGACCAAGTTCATGAGCCCTTGGAAAAAGG 938  
Db 2006 AGCAGCAGCAGAGATGAG 2065  
QY 939 TGGAGCGATCGAAAAACAACCCGCGCGGCGCCAAAGAGAGAGAGAGAGAGAGAGAGAG 998  
Db 2066 AGCAGCAGCAGATGAG 2125  
QY 999 ACGAAAGCAGTTCCCTGAGATCCGCAAGCAGCGCAGCTGACGAGCGCATGACAGCA 1058  
Db 2126 AGGATGAGCAG-----GAGCAGCAGAGATGACGACGACGAGATGACGACGAGAG 2179  
QY 1059 GGGTGGGCGAGCGGGGCACTGGGCTGTCCATGTCGCGCCCGCAGCCGACGACGAGAGTGT 1118  
Db 2180 ATGAGCAGCAGCAGAGATGAG 2239

QY 1119 CAGAGATCATCGATGCGCTCTCAGACGAGAGAACTTGAGAAAGATGCGCCAGCTGG 1178  
DB 2240 AGCAGAGACAGATGATGAACAGAGACAGACAGAGAGAGAGAGAGAGAGAGAGAG 2299  
QY 1179 CCGTATCCCGCCCATGCTGTACAGCGCTGACCAAGACGCGCATCAAGTTTCATCAATGA 1238  
DB 2300 AGAGAGTTAGAGAGACAG 2359  
QY 1239 AGGAGCTTATGCGCCGACCCCATGAAAGTGTATCAAAACCCCGCATGTCATGAAATGTGA 1298  
DB 2360 AGCAGAGAGTTAG 2419  
QY 1299 GTGAGCAG 1355  
DB 2420 AG 2479  
QY 1356 ---GCCGTATCGCATCATCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1412  
DB 2480 AGCAG 2539  
QY 1413 ACTGATCTAAGAAAGATGAACTATTAAGAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAG 1472  
DB 2540 AG 2599  
QY 1473 GCAAGAGCAG 1532  
DB 2600 AG 2659  
QY 1533 CCATGCCCCGACAG 1592  
DB 2660 AG 2719  
QY 1593 AGGAG 1652  
DB 2720 AAG 2779  
QY 1653 A 1653  
DB 2780 A 2780

RESULT 11  
US-09-894-273-1  
; Sequence 1, Application US/09894273  
; Patent No. 6756203  
; GENERAL INFORMATION:  
; APPLICANT: Kieff, Elliott D.  
; APPLICANT: Bailestas, Mary E.  
; APPLICANT: Kaye, Kenneth M.  
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO  
; TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE  
; FILE REFERENCE: 16412-10001R  
; CURRENT APPLICATION NUMBER: US/09/894,273  
; CURRENT FILING DATE: 2001-06-28  
; PRIOR APPLICATION NUMBER: US 60/109,422  
; PRIOR FILING DATE: 1998-11-19  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 1  
; LENGTH: 3489  
; TYPE: DNA  
; ORGANISM: Kapoli's sarcoma-associated herpesvirus  
US-09-894-273-1

Query Match 1.3%; Score 107.8; DB 3; Length 3489;  
Best Local Similarity 43.8%; Pred. No. 1.3e-10;  
Matches 631; Conservative 0; Mismatches 792; Indels 18; Gaps 3;

QY 219 AGCCGGGGAATGAAGGTCCCGAGAGCTCACTGTGCGCAGAGTCCCACTATACCTTCG 278  
DB 1352 AGCCACAGACAG 1411

QY 279 CCGAGCTGGGGAAGTCAAGATGAGAGTTCATTTGAAAGCAAGCCCTCGCTAGAGCTGC 338  
DB 1412 CACAGACAG 1471  
QY 339 TGCCTGACCCCTGCTGCTGCAAGCTGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 398  
DB 1472 AGCAG 1531  
QY 399 AAGACTTCAACAAAGAGCCGTAGCTGACGGGAGAGCTGGAACCGGAGTGTCTCCCCAGCC 458  
DB 1532 AGCAG 1585  
QY 459 CCCCAGACATGACCTGAGCTGAGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 518  
DB 1586 CACAG 1645  
QY 519 TCCAGAACATGAGACGCGCTGAGACCGAGAGATCAACATGTGAGAGAGAGAGATTTAAGC 578  
DB 1646 CACAG 1705  
QY 579 TGAAG 638  
DB 1706 CACAG 1765  
QY 639 CCGTGTCAACCGCCGCTCATGAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 698  
DB 1766 CACAG 1825  
QY 699 AGAACCGAG 758  
DB 1826 AGCAG 1885  
QY 759 AGCTGCGCTGTAACAACAG 818  
DB 1886 AGCAG 1945  
QY 819 ACCAG 878  
DB 1946 AGCAG 2005  
QY 879 AATGAG 938  
DB 2006 AGCAG 2065  
QY 939 TGAAG 998  
DB 2066 AGCAG 2125  
QY 999 AGGAAAG 1058  
DB 2126 AGGATGAG 2179  
QY 1059 GAGTGGAG 1118  
DB 2180 ATGAG 2239  
QY 1119 CAGAGATCATGATGAGCTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1178  
DB 2240 AGCAG 2299  
QY 1179 CCGTATCCCGCCCATGCTGTACAGCGCTGACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1238  
DB 2300 AGAGAGTTAG 2359  
QY 1239 AGGAGCTTATGCGCCGACCCCATGAAAGTGTATCAAAACCCCGCATGTCATGAAATGTGA 1298  
DB 2360 AGCAG 2419  
QY 1299 GTGAGCAG 1355  
DB 2420 AG 2479  
QY 1356 ---GCCGTATCGCATCATCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1412

Db 2480 AGCAGGACGAGGTTAGAGGAGCAGAGCAGAGTTAGAGGAGCAGGAGGTGGAAGAGC 2539  
Qy 1413 ACCTGACTAAGAAAGATGAACTATTAAGAGCTGTGTGAGACGAGACTATGAGCCCGCG 1472  
Db 2540 AAGACGAGAGGTGTAAGAGCAGAGCAGAGCAGAGAGAGCAAGATTAAGAGAGGTGG 2599  
Qy 1473 GCAGAGCCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1532  
Db 2600 AAGACGAGAGGAGCAGAGCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2659  
Qy 1533 CCATCCCCCGCAGCAGCCAGAGAGAGAAAGATGAGAGAGAGAGAGAGAGAGAGAGAG 1592  
Db 2660 AAGAGAGCAGAGGTTAGAGAGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2719  
Qy 1593 AGGAG 1652  
Db 2720 AAGACGAGAGAGAGAGAGAGAGTTAGAGAGGTGAGAGAGAGAGAGAGAGAGAGAG 2779  
Qy 1653 A 1653  
Db 2780 A 2780

RESULT 12  
US-08-770-379-20/c  
; Sequence 20, Application US/08770379  
; Patent No. 5849564  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Yuan  
; APPLICANT: Bohenzky, Roy A.  
; APPLICANT: Russo, James J.  
; APPLICANT: Edelman, Isidore S.  
; APPLICANT: Moore, Patrick S.  
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED  
; TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/770,379  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 52342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32207 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-770-379-20

Query Match 1.3%; Score 107.8; DB 2; Length 32207;  
Best Local Similarity 43.8%; Pred. No. 2.5e-10;  
Matches 631; Conservative 0; Mismatches 792; Indels 18; Gaps 3;

Qy 219 AGCCCGGAATGAACGATCCAGAGACTCAACTGCGGCGCAGAGTCCCATCTATATCTGCG 278  
Db 20645 AGCCACAGACAGAGAGAGCCAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20586  
Qy 279 CCGAGCTGGGAAATGACAGATGAGATCTTAAAGAAAGAGAGAGAGAGAGAGAGAGAGAG 338  
Db 20585 CACAGCAGCAGAGAGCAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20526  
Qy 339 TGCCCTGACCCCTGTGAGACCGATCCCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 398  
Db 20525 AGCAGCAG 20466  
Qy 399 AAGACTTACCAAG 458  
Db 20465 AGCAGAGAGCAG 20412  
Qy 459 CCCCAGACATGACCTGTGAG 518  
Db 20411 CACAGCAGCAGAGAGCAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20352  
Qy 519 TCAGAGATGAG 578  
Db 20351 CACAGCAGCAG 20292  
Qy 579 TGAG 638  
Db 20291 CACAGCAGCAG 20232  
Qy 639 CCGGTTCACCGCCCGCCATCCAGTGAAGCAGCCGAGCTGGTGAATCATCTTCAGAG 698  
Db 20231 CACAGCAGCAG 20172  
Qy 699 AGAACCAG 758  
Db 20171 AGCAGCAGAGAGATGAG 20112  
Qy 759 AGCTGCGCTGTGACCAAG 818  
Db 20111 AGCAGCAGAGAGATGAG 20052  
Qy 819 ACCAGGCGATCGAG 878  
Db 20051 AGCAGCAGCAG 19992  
Qy 879 AATGGAAG 938  
Db 19991 AGCAGCAGAGAGATGAG 19932  
Qy 939 TGAAGCGGATCGAAG 998  
Db 19931 AGCAGCAGAGAGATGAG 19872  
Qy 999 ACGAAAG 1058  
Db 19871 AGGATGAG 19818  
Qy 1059 GGGTGGGCGAG 1118  
Db 19817 ATGAGCAG 19758  
Qy 1119 CAGAGATCATGATGAG 1178  
Db 19757 AGCAGCAGAGAGATGAG 19698  
Qy 1179 CCGTGTATCCCGCCATGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1238  
Db 19697 AGGAGTTAG 19638  
Qy 1239 ACGGCTTATGAG 1298  
Db 19637 AGCAG 19578  
Qy 1299 GTGAGCAG 1355

Db 19577 AGGAGCAGAGGTTAGAGCAGAGCAGAGGAGTTAGAGCAGAGCAGAGTTAGAGG 19518  
QY 1356 ---GCTGATCGATCATCTTCTTGAGAGAGAGACAGTGTGCTGAGTCCCTCTATTACT 1412  
Db 19517 AGCAGAGCAGAGGTTAGAGAGCAGAGCAGAGGAGTTAGAGAGCAGAGAGTGAAGAGC 19458  
QY 1413 ACCTGACTAAGAGAACTATTAAGAGCTGTGTAAGCCGAGCTATCGCGCCGCG 1472  
Db 19457 AGAGCAGAGAGTGTGAAGAGCAGAGCAGAGCAGAGAGAGGAAATTAGAGAGGTGG 19398  
QY 1473 GCAAGAGCCAGCAGCAACAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1532  
Db 19397 AGGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGCAGAGTTAGAGAGGTGAAGAGC 19338  
QY 1533 CCATGCCCCCAGCAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1592  
Db 19337 AGGAGAGCAGAGGATTAGAGAGGTGAGAGAGCAGAGAGAGGTTAGAGAGAGGTGG 19278  
QY 1593 AGGAGAGAGAGAGAGCCGAGAGGTGAGAGAGCAGAGAGAGAGCTCTCAAGAGAGAGCAG 1652  
Db 19277 AGAGCAGAGAGCAGCAGAGGATTAGAGAGGTGAGAGAGCAGAGCAGAGAGGAGTGAAC 19218  
QY 1653 A 1653  
Db 19217 A 19217

RESULT 13  
US-08-757-669A-20/c  
Sequence 20, Application US/08757669A  
Patent No. 618751  
GENERAL INFORMATION:  
APPLICANT: Chang, Yuan  
APPLICANT: Bohenzky, Roy A.  
APPLICANT: Russo, James J.  
APPLICANT: Edelman, Isidore S.  
APPLICANT: Moore, Patrick S.  
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS  
TITLE OF INVENTION: SEQUENCES AND USES THEREOF  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/757,669A  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 45185-F  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32207 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-757-669A-20

Query Match 1.3%; Score 107.8; DB 3; Length 32207;  
Best Local Similarity 43.8%; Pred. No. 2.5e-10;  
Matches 631; Conservative 0; Mismatches 792; Indels 18; Gaps 3;  
QY 219 AGCCGGGAAATGAACGCTCCAGAGAGCTCCAGCTGGGCGAGAGTCCACTATCTTGC 278  
Db 20645 AGCCACAGCAGCAGGAGCCACAGCAGCAGAGAGCCCTGCGAGAGCCACACAGCAGAGC 20586  
QY 279 CCGAGCTGGGGAAATGAGAGATGAGATTCAATTGAAGCAAGCCGCTCGAGTGC 338  
Db 20585 CACAGCAGAGAGCCACAGCAGCAGAGAGCCCTGCGAGAGCCACACAGCAGAGCCAC 20526  
QY 339 TGCCTGACCCCTGCTGCGACCTGACCCCTGCTGCGACCGGCGAGCTGCGAGTCTG 398  
Db 20525 AGCAGCAGAGAGCCCTGCGAGAGCCACACAGCAGAGAGCCACACAGCAGAGCCACAGC 20466  
QY 399 AAGACTTACCAAGAACCGTAGCTTGAAGGCAAGTGGAACTGGATCTCCCCCAGCC 458  
Db 20465 AGCAGAGCCACAGCA-----GCAGAGCCACAGCAGAGAGGCCACAGCAGCAGAGC 20412  
QY 459 CCCCAGCACTGACCTGAGCTGAGCTGTGCTGCGCGCAGCGGCTGTCAGAGAGAGCTGA 518  
Db 20411 CACAGCAGAGAGCCACAGCAGCAGAGAGCCACAGCAGAGCCACAGCAGAGAGCCACAGCAGAGAGC 20352  
QY 519 TCCAGAACTGGAACCGCTGGAACCGAGATCACTATGTTAGAGCAGCAGAGTCTTAAGC 578  
Db 20351 CACAGCAGCGGAGCCACAGCAGCAGAGAGCCACAGCAGCGGAGCCACAGCAGCGGAGC 20292  
QY 579 TGAAGAGAGAGCAGCAACAGCTGGAAGAGAGAGCTCCAGAGCCGCGGAGCTGAAGAGC 638  
Db 20291 CACAGCAGCGGAGCCACAGCAGCAGAGAGCCACAGCAGCGGAGCCACAGCAGCGGAGC 20232  
QY 639 CCGTGTCAACCGCCGCCCATCTGAGTGAAGCAGCAGCCTGTGTGAGATCATTAAGAGC 698  
Db 20231 CACAGCAGCGGAGCCACAGCAGCAGAGATGAGCAGAGAGTGAAGAGCAGCAGCAGAGT 20172  
QY 699 AGAACCGAAGAGAGCTGAAGCTGACATCGATTCTGGAAGGCTCGGGGCCCGCAGGTGG 758  
Db 20171 AGCAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGCAGCAGATGAGCAGCAGCAGAGT 20112  
QY 759 AGCTGCGCTGTACAAACAGCCCTCCGACACCCGGAGATCATGAACATCAAAATTA 818  
Db 20111 AGCAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGAGT 20052  
QY 819 ACCAGGCAATGCGAAGAGCTTAATTTGTAATTCAAGAGAGAGATCAAGCTCGGAAGC 878  
Db 20051 AGCAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGCAGAGATGAGCAGCAGCAGAGT 19992  
QY 879 AATGGAAGCAGAACTTCTGCGAGCTATGACCAAGCTCATGAGAGCCTTGAAGAAAAAG 938  
Db 19991 AGCAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGAGT 19932  
QY 939 TGGAGGCAATCGAAACAAACCGCGCGCGGCGGCGCAAGAGAGCAAGTGGCGAGTACT 998  
Db 19931 AGCAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGCAGCAGCAGATGAGCAGCAGAG 19872  
QY 999 AGGAAAGCAGTTCCCTGAGATCCGACAGCAGCGCAGCTGCGAGAGCCGATCAAGACA 1058  
Db 19871 AGGATGAGCAG-----GACAGCAGAGATGAGCAGCAGAGATGAGCAGCAGCAGCAG 19818  
QY 1059 GGGTGGGCGCAGCGGGCAGTGGGCTGTCCATGTGCGCCCGCCAGCAGCAGAGAGTGT 1118  
Db 19817 ATGAGCAGCAGCAGAGATGAGCAGCAGCAGCAGCAGAGATGAGCAGCAGCAGAGTGA 19758  
QY 1119 CAGAGATCATGATGCGCTTCAAGCAGAGAGAACTTGAAGAAAGCAGATGCGCAGCTGG 1178  
Db 19757 AGCAGCAGCAGATGAAACAGAGCAGCAGAGAGAGAGAGCAGCAGAGAGCAGAGAGC 19698  
QY 1179 CCGTATCCCGCCCATGCTGTGAGAGCTGACAGCAGCAGAGCAGATCAAGTTCAATCA 1238  
Db 19697 AGGAGTTAGAGAGCAGAGCAGAGGATTAGAGATTCAGAGCAGAGGATTAGAGAGCAG 19638  
QY 1239 ACGGGCTATGAGCCGACCCCATGAAAGGTGACAAAGACCGCAGGTCATGAACATGTGA 1298

Db 19637 AGCAGAGTTAAGAGACGAGACGAGATTAGAGAGCAGAGCAGAGATTAGAGAGC 19578  
QY 1299 GTGACAGAGAGAGACCTTCCGGAGAAAGTTCAATGACGATCCCAAGAACTTTG--- 1355  
Db 19577 AGGAGCAGAGATTAGAGAGCAGAGCAGAGATTAGAGAGCAGAGAGATTAGAG 19518  
QY 1356 ---GCCCTGATCGATCATCTTCTTGAGAGAGAGACAGTGGCTGATGTCCTCTTATTACT 1412  
Db 19517 AGCAGAGACGAGATTAGAGAGCAGAGCAGAGATTAGAGAGCAGAGATTAGAGAGC 19458  
QY 1413 ACCTGACTAAGAAAGATGAACTAATAGAGCTGCTGTGAGAGGCTATTCGGCCCGG 1472  
Db 19457 AAGACAGAGAGGTGAGAGCAAGAGCAGAGCAGAGAGCAGAGATTAGAGAGGTGG 19398  
QY 1473 GCAGAGCCAGAGCAACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1532  
Db 19397 AGGAGCAGAGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 19338  
QY 1533 CCATGCCCCGACGACGACGAGAGAGAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGC 1592  
Db 19337 AGGAGAGCAGAGATTAGAGAGGTGGAAAGCAGAGAGCAGAGAGATTAGAGAGGTGG 19278  
QY 1593 AGGAGAGAGAGAGAGCCGAGGTGGAGAAAGCAAGAGAGAGCCTCTCAAGAGAGAGAG 1652  
Db 19277 AAGAGCAGAGAGCAGAGAGATTAGAGAGGTGAAAGAGCAGAGAGCAGAGAGGTGAAAC 19218  
QY 1653 A 1653  
Db 19217 A 19217

RESULT 14  
US-09-230-371A-20/c  
; Sequence 20, Application US/09230371A  
; Patent No. 6348586  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Yuan  
; APPLICANT: Bohenzky, Roy A  
; APPLICANT: Russo, James J  
; APPLICANT: Edelman, Reidore S  
; APPLICANT: Moore, Patrick S  
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: 45185-G-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/230,371A  
; CURRENT FILING DATE: 1999-11-17  
; PRIOR APPLICATION NUMBER: PCT/US97/13346  
; PRIOR FILING DATE: 1997-07-22  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 32207  
; TYPE: DNA  
; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-230-371A-20

Query Match 1.3%; Score 107.8; DB 3; Length 32207;  
Best Local Similarity 43.8%; Pred. No. 2.5e-10;  
Matches 631; Conservative 0; Mismatches 792; Indels 18; Gaps 3;  
QY 219 AGCCCGGAGATGAACGCTCCAGAGACTCCAGCTGCGCCAGAGTCCCATATCTTGC 278  
Db 20645 AGCCACAGACAGAGAGCCACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20586  
QY 279 CCGAGCTGGAGAGTCAAGATGAGATTATTGAAAGCAAGGCCCTCGGCTGAGAGCTGC 338  
Db 20585 CACACAGAGAGAGAGCCACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20526  
QY 339 TGCCTGACCCCTGCTGCGAGCGTCAACCTGCTGGCCAGCGGAGAGAGCTGAGATCTG 398  
Db 20525 AGCAGCAG 20466

QY 399 AAGACCTCAACCAAGAGACCGTACCTGAGCGGCAAGCTGGAACCGGTGTCTCCCCCAGCC 458  
Db 20465 AGCAGAGAGCCACAGCA-----GCAGAGCCACAGACAGAGAGCCACAGCAGCAGAGC 20412  
QY 459 CCCCAGCAGTGAACCTGAGCTGAGCTGTGTCGCCACCGCTGTCCAAAGAGAGAGTGA 518  
Db 20411 CACAGCAGAGAGCCACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20352  
QY 519 TCCAGAACATGAGCCGCGTGAACCGAGAGATCAACATGTAAGAGCAGAGATCTTAAAGC 578  
Db 20351 CACAGCAGCGGAGGCCACAGCAGCGGAGAGCCACAGCAGCGGAGGCCACAGCAGCGGAGC 20292  
QY 579 TGAAGAAAGCAGCAACAGCTGAGAGAGAGAGCTGCCAACCCGCTCCAGCTTGAAAGC 638  
Db 20291 CACAGCAGCGGAGGCCACAGCAGCGGAGGCCACAGCAGCGGAGGCCACAGCAGCGGAGC 20232  
QY 639 CCGTGTCAACCGCGCCCATCGATCGATGAAGCAACCGCAGCTGGTGAATCATCTACAGC 698  
Db 20231 CACAGCAGCGGAGGCCACAGCAGCGGAGGCCACAGCAGCGGAGGCCACAGCAGCGGAGC 20172  
QY 699 AAGACCGGAGAGAGCTGAAGCTGCAATCGGATTTCTGAAAGCCCTGGAGCCCAAGTGG 758  
Db 20171 AGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGATG 20112  
QY 759 AGCTGCCCTGTACAACCAAGCCTTCGACACCCGCGAGTATCATAGAGAAATCAATAATA 818  
Db 20111 AGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGATG 20052  
QY 819 ACCAGCGCATCGGAGAGAGCTAATCTTGTACTTCAAGAGAGAAATCAGCTCGGAAGC 878  
Db 20051 AGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGATG 19992  
QY 879 AATGAGCAGAGAGTTCTGCAAGCTGATGACCACTCATGAGGCTTGAAAAAAGG 938  
Db 19991 AGCAGCAGCAGAGATGAGCAGCAGCAGATGAGCAGCAGCAGAGATGAGCAGCAGATG 19932  
QY 939 TGGAGCGATGAGAAACCAACCCCGCCGCGGCCAAGAGAGCAAGGTGCGCGAGTACT 998  
Db 19931 AGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGCAGC 19872  
QY 999 AGGAAAGCAGATCTCCCTGAGATCCGCAAGCAGCGGAGCTGCAAGAGAGCATGAGAGCA 1058  
Db 19871 AGGATGAGCAG-----GAGCAGCAGAGTGAAGCAGCAGCAGAGATGAGCAGCAGAG 19818  
QY 1059 GGGTGGGCGACGCGGCGAGTGGGCTGTCTATGTGGCGCCGCGAGCAGAGCATGAGTGT 1118  
Db 19817 ATGAGCAGCAGAGCAGAGATGAGCAGCAGCAGAGATGAGCAGCAGAGATGAGCAGAGATGAGC 19758  
QY 1119 CAGAGATCATGATGAGCTCTCAAGCAGAGAGAACTTGAAGAGAGATGCGCCAGCTGG 1178  
Db 19757 AGCAGCAGCAGAGATGAACAGAGCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 19698  
QY 1179 CCGTATCCCGCCCATGCTGTATGAGACGCTGAACAGCAGCGCATCAAGTTCAATGAATGA 1238  
Db 19697 AGGAGTTAGAGAGAGAGAGAGAGAGAGAGAGATGAGAGATCAGAGACAGAGATTAGAGAGAGC 19618  
QY 1239 ACGGCTTATGAGCCGACCCCATGAGGTGTAACAAGCCGCAAGCTCATGAACATGTGA 1298  
Db 19637 AGCAGAGATTAGAGAGAGAGAGAGAGAGATTAGAGAGCAGAGAGAGAGAGAGAGAGC 19578  
QY 1299 GTGAGCAGAGAGAGAGACCTTCCGGAGAAATTATGACAGTATCCCAAGAACTTTG--- 1355  
Db 19577 AGGAGCAGAGAGATTAG 19518  
QY 1356 ---GCCCTGATCGATCATCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1412  
Db 19517 AGCAGAGCAGAGATTAG 19458  
QY 1413 ACCTGACTAAGAAAGATGAACTAATAGAGCTGCTGTGAGAGGCTATTCGGCCCGG 1472  
Db 19457 AAGACAGAGAGGTGAGAGCAAGAGCAGAGCAGAGAGCAGAGATTAGAGAGGTGG 19398  
QY 1473 GCAGAGCCAGAGCAACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1532  
Db 19397 AGGAGCAGAGAGAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGC 19338  
QY 1533 CCATGCCCCGACGACGACGAGAGAGAAAGATGAGAGAGAGAGAGAGAGAGAGAGAGC 1592  
Db 19337 AGGAGAGCAGAGATTAGAGAGGTGGAAAGCAGAGAGCAGAGAGATTAGAGAGGTGG 19278  
QY 1593 AGGAGAGAGAGAGAGCCGAGGTGGAGAAAGCAAGAGAGAGCCTCTCAAGAGAGAGAG 1652  
Db 19277 AAGAGCAGAGAGCAGAGAGATTAGAGAGGTGAAAGAGCAGAGAGCAGAGAGGTGAAAC 19218  
QY 1653 A 1653  
Db 19217 A 19217





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 12, 2006, 05:34:50 ; Search time 27353 Seconds  
(without alignments)  
17811.765 Million cell updates/sec

Title: US-09-522-753-4

Perfect score: 8564.8  
Sequence: 1 catctgcggcctccacacagc.....caaaaaaaaaaaaaaaaaa 8571

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Genemb1:\*  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_dr:\*  
9: gb\_ro:\*  
10: gb\_sts:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_hcg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8541	99.7	8561	8 AF113003	AF113003 Homo sapi
2	8324.6	97.2	8686	8 AF125672	AF125672 Homo sapi
3	8226.6	96.1	9053	6 AR447713	AR447713 Sequence
4	8038.6	93.9	8548	8 AY965853	AY965853 Homo sapi
5	5455	63.7	5989	8 HSU37146	U37146 Human sllen
6	5037.2	58.8	8544	9 AF125671	AF125671 Mus muscu
7	4981.6	58.2	8388	9 AF113001	AF113001 Mus muscu
8	4578.2	51.1	7465	9 AF113002	AF113002 Mus muscu
9	4253.8	49.7	4686	8 AB209089	AB209089 Homo sapi
10	3963.8	46.3	6339	6 CQ722208	CQ722208 Sequence
11	2759.8	32.2	2842	8 BC004326	BC004326 Homo sapi
12	2618.6	30.6	2930	8 S83390	S83390 T3 receptor
13	1703.4	19.9	2964	9 BC047524	BC047524 Mus muscu
14	1132.4	13.2	205283	8 AC073916	AC073916 Homo sapi
15	948.4	11.1	161970	14 AC027706	AC027706 Homo sapi
16	898.4	10.5	956	6 BD270289	BD270289 Human nuc
17	898.4	10.5	956	6 AX035226	AX035226 Sequence
18	898.4	10.5	956	6 AX342080	AX342080 Sequence

19	898.4	10.5	956	6 AX342140	AX342140 Sequence
20	898.4	10.5	956	6 AX441143	AX441143 Sequence
21	856.4	10.0	1907	5 CR926317	CR926317 Xenopus t
22	854.6	10.0	1896	5 BC082706	BC082706 Xenopus 1
23	854.6	10.0	1917	5 BC054296	BC054296 Xenopus 1
24	732.8	8.6	79494	14 AC068837	AC068837 Homo sapi
25	652.8	7.6	752	6 CQ769363	CQ769363 Sequence
26	650	7.6	650	6 AX677743	AX677743 Sequence
27	631.6	7.4	7949	8 AB028970	AB028970 Homo sapi
28	630	7.4	7940	6 CQ847916	CQ847916 Sequence
29	630	7.4	7940	6 AX578069	AX578069 Sequence
30	630	7.4	7940	6 AF044209	AF044209 Homo sapi
31	629.6	7.4	7780	9 MMUJ5312	U35312 Mus musculu
32	626.8	7.3	3120	8 AF303586	AF303586 Homo sapi
33	626.8	7.3	8018	8 AF087856	AF087856 Homo sapi
34	615.8	7.2	7254	9 BC086657	BC086657 Mus muscu
35	572.2	6.7	8626	5 BC099620	BC099620 Xenopus t
36	570	6.7	725	10 BV448749	BV448749 NCOR2_795
37	568.4	6.6	8959	5 AF495886	AF495886 Xenopus 1
38	555	6.5	555	6 AX677866	AX677866 Sequence
39	554.8	6.5	560	8 HSU80761	U80761 Homo sapien
40	553.6	6.5	560	8 HSU80750	U80750 Homo sapien
41	542.4	6.3	718	6 AX753058	AX753058 Sequence
42	540.6	6.3	1891	8 BC050594	BC050594 Homo sapi
43	534	6.2	534	6 AX778323	AX778323 Sequence
44	533.4	6.2	1850	8 BC058511	BC058511 Homo sapi
45	533.2	6.2	1808	8 BC068996	BC068996 Homo sapi

## ALIGNMENTS

RESULT 1	AF113003	8561 bp	mRNA	linear	PRI 20-MAR-1999
LOCUS	AF113003				
DEFINITION	Homo sapiens silencing mediator of retinoic acid and thyroid hormone receptor alpha mRNA, complete cds.				
ACCESSION	AF113003				
VERSION	AF113003.1	GI:4454551			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 8561)				
AUTHORS	Ordentlich,P., Downes,M., Xie,W., Genin,A., Spinner,N.B. and Evans,R.M.				
TITLE	Unique forms of human and mouse nuclear receptor corepressor SMRT				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (6), 2639-2644 (1999)				
PubMed	10077563				
REFERENCE	2 (bases 1 to 8561)				
AUTHORS	Downes,M.R., Ordentlich,P. and Evans,R.M.				
TITLE	Direct Submission				
JOURNAL	Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA				
FEATURES	Location/Qualifiers				
source	1..8561				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/chromosome="12"				
	/map="12q24"				
	/cfeature_type="pituitary"				
	2..7555				
	/function="transcriptional co-repressor"				
	/note="bSMRT alpha; longer isoform than previously reported"				
	/codon_start=1				
	/product="silencing mediator of retinoic acid and thyroid hormone receptor alpha"				
	/protein_id="AAD20946.1"				

Search History

/db\_xref="GI:4454552"  
/translation="MSGTQVLAQWTRATERYPPHSLSYPIVQIARTHTDVLGLYOH  
HSDRYASHLSIPSIIOQORRRSLLSBPNGMERQELHRLRESHSYLPRLKSMER  
IESKRRLLELPDLRPLATGTGPASBDLTXDRSLTGLRLEVPSPRPTDPEL  
ELVPRLSKEELIOMNDVDRBITWVQOISLTKKKOOLJEEBAKPPPEPVBP  
IESKRSLVOIITYENRKKKAEAAHRLIEGLGQVELPLVRSNDROYHNMKIYOM  
BKLLYPRBHARKKOKKFCORYDOLMBALKEKVRLENNPRAHESVYRYEAM  
KOPPEIRKORELOERMOSHVGORGSGLSNAARSEHVSSEITDGLSEONLKKOQL  
AVIPMLYDADQORIKFINMGMADPMKVYDRQVMNMSBQESKEFEKEMOHPN  
FOLIASFLERKIVACVLYLTLKNENKSLVRSYRGRGSOOQOQOQOQOQOQOQ  
OQPMRPSQOEKDEKEKEKEBEKEPEVNDKEDLKEKDTDSGENDENAVAS  
KGRKTASQGRKRLITREMANEASBEALITPOQASKEIASMLNSRTEEMERAK  
KGLLHGRMSAIAMVSGKTVSOCKNIFRYNKKRONDELIIQOKHKLMEKERNARK  
KKGAAPAASEBAFPPEVEDEMEASGVSENEWESEBEALHASGENVREKESPA  
TVNNSVDESISPHTEAKDQNGQPPATLGADGPPPTPRRTSRAPRIEPT  
KAEBAKCTPPAPPSAPPPVUPVEKEKEBEALAAPEVEBEKOPPAEELADTG  
KLEBPVSECTEEAEEGPAKQABAEALTAEGALKAKEKGGSGATATKSGAPOD  
SSSASGADVEDBAEGDKNLLSPRSLITPTGDPNANASPOPDILKOIKORAAA  
IPRIQVTKYHREBDAATKAPAPRPPQMLQPSBAPQOQSSPRKESRPARPA  
DKEATPAEQKLPEDPCWTSGLPFPVPREVITKASPHAPDSATYAPRPHLPGL  
HDTPAPVLRPPTISNPRLISSAKHPSVLERQIGAIISQMSVOLHVPSEHAKAVG  
PVTMGLPLMPDKKLAPFSGVQEOQLSPRGAQBPBSLGVPTAOEASVLRGTALGSP  
GSGITKGIJSTVPDSAITYRGSITHGPDVLYKGTITRIIGDSPERLDRGSDS  
LPGHVIYEGKKGVLSYEGMSVTCSEKEDRSSGPHETAAKRTYDMMEGVGR  
AISASISGIMGRALIPREHSPHILKEOHITRGSITQGPBSYVAQEDVYLRBEKLL  
KREGTPPPPSPRSDLTAEKTYDLEPLKLPAHBEGLVATVKAGSIEHITPREBLHT  
PELPAPRPLKESITQGTPLKYDTGASTGSKHDVRLISSPRTPEVPHLPVMA  
DHALERACYEBSLKSPEGTASSSGSILRGAIVIPREIAGKROSLTYEDGAPAG  
HPRGSVPTMRPTPRLQESLSSKASODRLKTSTPRILASPHSYVEHHPHPSIP  
YEHILRGVSGVDLYRSHPLAFDPPTSPRGIPLDAAAYLRRHLAHPNTYHLLPPY  
LIRKPDPALENRQTIINDYITTSQOMHNNTATMAQRPDMRGSLSPRESSALNTAA  
GREGIIDSQVPHLPVPPTGTPYAMDRLAYLPLAPPSRSHSSPLSPGATH  
LKPPTTSSSERERDRERDREREKSLITSTTVBERAPIMRGETOSSSSSG  
GGGSSSRPASHAHQHSIPRTOQALQORPSVLNMGGIIITAAVEPSPTYLRS  
TSSGVPAPATTPPATHCPLGLDGVPTLMEPVLYLPKAPARVAPRPAUDGSP  
FLAKPARGLEPASPSPKSGSPRPVLPVSGHATITARPALHA.PHNASPPAPPA  
SASDHRKCTGKSPSIOELERSLGYNHSSSPGVEVSSPSLTHDGLKXHL  
BELDLSHLEBELPKOPFVKUGSEAAHLPILPLPSPSSPLQTPAGVKGORV  
VTLAHLISEVITYDYTRHHPQOLSAPLPAPLSFQASGPVLDLRRPBDLPPDH  
GAPARGSPHSEGKRSPEBNKTSVLGGGEGDIEVPSPGMEHPPGHSRAVPLRYD  
GQOTSPRWKSGSPNTSOPAPFSPKLTBSNAGVMSKQOINKKLTNNRREPRNT  
SOPGTEIFNMPAITGTGMLTWSOAVOEASTNMGLI.IIIRALMGKYOMESPLS  
ANAPNINASASLPAAMPRTIADGSDHILTPSGGGRAKXVGRSSSRKASPAEGA  
SDRPPSVSVSEBDCKNRTPLTNRVMDRPSAGSTPFPPNPLMLRQAVMASPP  
PGLPAGSGPLAGPHHAMDPEBKPLCSQYETLSDSE"

ORIGIN

Query Match	99.7%	Score 8541	DB 8	Length 8561
Best Local Similarity	99.9%	Pred. No. 0		
Matches 8560	Conservative 1	Mismatches 0	Indels 10	Gaps 1
QY	1	CATGTCGGGCTCCACA	CAGCTTGTGGCACA	GAGGAGGCACTGAGCCCGCTAACCC 60
DB	1	CATGTCGGGCTCCACA	CAGCTTGTGGCACA	GAGGAGGCACTGAGCCCGCTAACCC 60
QY	61	GCCCAACAGCTTTTCT	TCTACCCAGTGCAGAT	CGCCGGAGCGCAACGGAAGCTGGGCTCTT 120
DB	61	GCCCAACAGCTTTTCT	TCTACCCAGTGCAGAT	CGCCGGAGCGCAACGGAAGCTGGGCTCTT 120
QY	121	GGAAGTACAGACCACT	CCCGGAGCTATGCCCT	CCAGCTGGCGGGCTTCATATCA 180
DB	121	GGAAGTACAGACCACT	CCCGGAGCTATGCCCT	CCAGCTGGCGGGCTTCATATCA 180
QY	181	GCCCAAGCGGCGGAGG	CCCTCCCTGCTGTCTG	ATGATTCAGCCCGGGAATGAACGCTCCA 240
DB	181	GCCCAAGCGGCGGAGG	CCCTCCCTGCTGTCTG	ATGATTCAGCCCGGGAATGAACGCTCCA 240
QY	241	GGAAGTCCACCTGCGG	CCAGAGTCCCATCTCA	TCTGCGGAGCTGGGAAAGTCAAGAT 300
DB	241	GGAAGTCCACCTGCGG	CCAGAGTCCCATCTCA	TCTGCGGAGCTGGGAAAGTCAAGAT 300
QY	301	GGAAGTTCAATTGAAAG	GAAGCGCCCTTGCGT	AGAGCTGCTGACCCCTGCTGAGACC 360
DB	301	GGAAGTTCAATTGAAAG	GAAGCGCCCTTGCGT	AGAGCTGCTGACCCCTGCTGAGACC 360

DB	301	GGAAGTTCAATTGAAAG	GAAGCGCCCTTGCGT	AGAGCTGCTGACCCCTGCTGAGACC 360
QY	361	GTACACCCCTGTGCTG	GCACCGGCGCAAGCT	CTGGGGAATCTGAAGCTTCCACGAAGACCTTAG 420
DB	361	GTACACCCCTGTGCTG	GCACCGGCGCAAGCT	CTGGGGAATCTGAAGCTTCCACGAAGACCTTAG 420
QY	421	CTTGAACGGGCAAGCT	GGAAACCGGTGTC	CCCCCGGCGCAACATGACCCCTGAGCT 480
DB	421	CTTGAACGGGCAAGCT	GGAAACCGGTGTC	CCCCCGGCGCAACATGACCCCTGAGCT 480
QY	481	GGAAGTGTGCGCGGCA	CGGCTGTCCAAAGAG	AGAGGTGATCCAGAACATGACCGGCTGGA 540
DB	481	GGAAGTGTGCGCGGCA	CGGCTGTCCAAAGAG	AGAGGTGATCCAGAACATGACCGGCTGGA 540
QY	541	CCGAGAGATCAACAT	GTGTAGACAGACAAT	CTCTAAGCTGAAGAAAGACAGCAACAGCT 600
DB	541	CCGAGAGATCAACAT	GTGTAGACAGACAAT	CTCTAAGCTGAAGAAAGACAGCAACAGCT 600
QY	601	GGAAGAGAGGCTGCG	CAAGCCGCCCGAGC	CTTGAGAAAGCCCTGTGACCGCGGCCCATGGA 660
DB	601	GGAAGAGAGGCTGCG	CAAGCCGCCCGAGC	CTTGAGAAAGCCCTGTGACCGCGGCCCATGGA 660
QY	661	GTGGAAGCACCGGAC	CTGTGTGATCATAT	CAACAGAGAAACCGGAAAGGCTGAAGC 720
DB	661	GTGGAAGCACCGGAC	CTGTGTGATCATAT	CAACAGAGAAACCGGAAAGGCTGAAGC 720
QY	721	TGCACATCGGATTTG	GAAGGCTTGAGG	CCCGAGCTGCGGTGTACACCAAGCC 780
DB	721	TGCACATCGGATTTG	GAAGGCTTGAGG	CCCGAGCTGCGGTGTACACCAAGCC 780
QY	781	CTCCGACACCCCGG	CAGTATCATGAGA	CAATCAATAATTAACGAGGGATGCGGAAGAGCT 840
DB	781	CTCCGACACCCCGG	CAGTATCATGAGA	CAATCAATAATTAACGAGGGATGCGGAAGAGCT 840
QY	841	AATCTTGAATCTTCA	GAGAGGAAATCAAG	CTGGAAGCAATGGAAGCAAGATTCTGCA 900
DB	841	AATCTTGAATCTTCA	GAGAGGAAATCAAG	CTGGAAGCAATGGAAGCAAGATTCTGCA 900
QY	901	GCGCTATGACCAAG	CTCATGAGAGGCT	TGTGAAAAAAGGTGAGCGCATGAAAAACAACC 960
DB	901	GCGCTATGACCAAG	CTCATGAGAGGCT	TGTGAAAAAAGGTGAGCGCATGAAAAACAACC 960
QY	961	GCGCGCGCGGCGCA	AGAGAGCAAGGTG	CGCGCAATCTGAGAT 1020
DB	961	GCGCGCGCGGCGCA	AGAGAGCAAGGTG	CGCGCAATCTGAGAT 1020
QY	1021	CCGCAAGCAGCGGAG	CTGTGACAGAGCG	CATGACAGAGGAGGAGCGGAGGAGGAGTGG 1080
DB	1021	CCGCAAGCAGCGGAG	CTGTGACAGAGCG	CATGACAGAGGAGGAGGAGGAGGAGTGG 1080
QY	1081	GCTGTCCATGTTGAG	CCCGCCGAGGAG	CGAGAGTGTCAAGATCATGATGAGCTCTTC 1140
DB	1081	GCTGTCCATGTTGAG	CCCGCCGAGGAG	CGAGAGTGTCAAGATCATGATGAGCTCTTC 1140
QY	1141	AGAAGCAGAGAACT	TGAGAGAGCAATG	CGCTGGCCGTGATCCCGGCCCATGCTGTGA 1200
DB	1141	AGAAGCAGAGAACT	TGAGAGAGCAATG	CGCTGGCCGTGATCCCGGCCCATGCTGTGA 1200
QY	1201	GGAAGTGTGACCAAG	CTGATCAAGATG	CAATGAACGGGCTTATGAGCGGACCCCAT 1260
DB	1201	GGAAGTGTGACCAAG	CTGATCAAGATG	CAATGAACGGGCTTATGAGCGGACCCCAT 1260
QY	1261	GGAAGTGTGACCAAG	CTGATCAAGATG	CAATGAACGGGCTTATGAGCGGACCCCAT 1320
DB	1261	GGAAGTGTGACCAAG	CTGATCAAGATG	CAATGAACGGGCTTATGAGCGGACCCCAT 1320
QY	1321	CGGAGGAAAGTTCA	TGACAGCATCCCA	GAACCTTTGGCTGATCGATCTCTTCTGAGAG 1380
DB	1321	CGGAGGAAAGTTCA	TGACAGCATCCCA	GAACCTTTGGCTGATCGATCTCTTCTGAGAG 1380
QY	1381	GGAAGCAGTGTGAG	TGTGCTCTTATTA	CTAAGCTGAAGAAAGTGAAGATCTATA 1440
DB	1381	GGAAGCAGTGTGAG	TGTGCTCTTATTA	CTAAGCTGAAGAAAGTGAAGATCTATA 1440

[illegible]

QY	2511	GGAGGAGCAAGAGCCCCCGCGGCTTGAAGACTGGCACTGTGACACAAAGAAAGCCCGAGA	2580
Db	2521	GGAGGAGCAAGAGCCCCCGCGGCTTGAAGACTGGCACTGTGACACAAAGAAAGCCCGAGA	2580
QY	2581	GCCCCCTCAAGAGCGAGTGTCACGAGAGAAAGCCGAGAGAGGGGCGCGCCAAAGGGCAAGACGC	2640
Db	2581	GCCCCCTCAAGAGCGAGTGTCACGAGAGAAAGCCGAGAGAGGGGCGCGCCAAAGGGCAAGACGC	2640
QY	2641	GGAGGCGCTGAGGCGCACGCGCCGAGGGGGCGCTCAAGGCAGAGAAAGAGAGGGCGGGAG	2700
Db	2641	GGAGGCGCTGAGGCGCACGCGCCGAGGGGGCGCTCAAGGCAGAGAAAGAGAGGGCGGGAG	2700
QY	2701	CGGCAAGGGCCACCACTGCGCAAGAGCTGGGGGCGCCCCCAGGACAGCGCACTTCAATGTCTAC	2760
Db	2701	CGGCAAGGGCCACCACTGCGCAAGAGCTGGGGGCGCCCCCAGGACAGCGCACTTCAATGTCTAC	2760
QY	2761	CTGCAAGTACAAACGAGGTGATGTGAGGGCCGAGGGGGGGGACAAAGAACGGGCTGTCTCCC	2820
Db	2761	CTGCAAGTACAAACGAGGTGATGTGAGGGCCGAGGGGGGGGACAAAGAACGGGCTGTCTCCC	2820
QY	2821	AAGGCGCCAGCTCTCTCAACCCCGACTGGCGAACCOCGGGGCCAAATGCTCAACCCCAAGAGCC	2880
Db	2821	AAGGCGCCAGCTCTCTCAACCCCGACTGGCGAACCOCGGGGCCAAATGCTCAACCCCAAGAGCC	2880
QY	2881	ACTGAGCCTGAAGACGCTGAAGACGAGCGAGCGGCTGCATCTCCGCCCATTCGAGGTACCAA	2940
Db	2881	ACTGAGCCTGAAGACGCTGAAGACGAGCGAGCGGCTGCATCTCCGCCCATTCGAGGTACCAA	2940
QY	2941	AGTCCATGAGCCCCCGGGAGGAGCGAGCTCCCAACAAAGCAGCTCCCCCAGGCCCAAC	3000
Db	2941	AGTCCATGAGCCCCCGGGAGGAGCGAGCTCCCAACAAAGCAGCTCCCCCAGGCCCAAC	3000
QY	3001	GCACACCGGAAACCTGCGACGCGGAGAGCGAAGCCCTCAGACGAGCTGGACAGACGCCCG	3060
Db	3001	GCACACCGGAAACCTGCGACGCGGAGAGCGAAGCCCTCAGACGAGCTGGACAGACGCCCG	3060
QY	3061	GGGCAAGAGCAGAGAGCCCGGACCCCGCGCGACAAAGAGGGCTTGGCAGCCGAGGCCCA	3120
Db	3061	GGGCAAGAGCAGAGAGCCCGGACCCCGCGCGACAAAGAGGGCTTGGCAGCCGAGGCCCA	3120
QY	3121	GAACTGCTGGGGAGCCCCCTTGTCTGGACTTCCGGGCTGTGCTTCCCCCGTCCCCCG	3180
Db	3121	GAACTGCTGGGGAGCCCCCTTGTCTGGACTTCCGGGCTGTGCTTCCCCCGTCCCCCG	3180
QY	3181	TGAGGTGATCAAGGCGTCCCCCGCAATGGCCCGGAAACCCCTCAAGCTTCTCTAACGCTCCACC	3240
Db	3181	TGAGGTGATCAAGGCGTCCCCCGCAATGGCCCGGAAACCCCTCAAGCTTCTCTAACGCTCCACC	3240
QY	3241	TGCTCACCACTGCGCCCTGGGCGCTGCATGACACTGCGCGGCGCGTCTCTGCGCGGCCACC	3300
Db	3241	TGCTCACCACTGCGCCCTGGGCGCTGCATGACACTGCGCGGCGCGTCTCTGCGCGGCCACC	3300
QY	3301	CACCACTTCCAAACCGCGCTTCCCTCATCTCTCTGCGAAGACCCCAAGCGTCTCTGAGAG	3360
Db	3301	CACCACTTCCAAACCGCGCTTCCCTCATCTCTCTGCGAAGACCCCAAGCGTCTCTGAGAG	3360
QY	3361	GCAATATGAGTGCATCTTCCCAAGGAATGTGAGTGCACACTCAAGTCCCGTAACTCAGAGCA	3420
Db	3361	GCAATATGAGTGCATCTTCCCAAGGAATGTGAGTGCACACTCAAGTCCCGTAACTCAGAGCA	3420
QY	3421	TGCCAAGGCCCCGGTGGGCGCTGTGCACATGGGGCTGCCCTGCGCAATGACCCCAAAA	3480
Db	3421	TGCCAAGGCCCCGGTGGGCGCTGTGCACATGGGGCTGCCCTGCGCAATGACCCCAAAA	3480
QY	3481	GCTGGCACCTTTCAGCGGAGTGAAGACGAGCAGCTGTCCCAACGGGGGCAAGCTTGGGGC	3540
Db	3481	GCTGGCACCTTTCAGCGGAGTGAAGACGAGCAGCTGTCCCAACGGGGGCAAGCTTGGGGC	3540
QY	3541	ACCGGAGAGCCTGGGGGTGCCCAAGGCCCAAGAGGCGTCCGTGTGAGAGGAGACAGCTT	3600
Db	3541	ACCGGAGAGCCTGGGGGTGCCCAAGGCCCAAGAGGCGTCCGTGTGAGAGGAGACAGCTT	3600
QY	3601	GGGCTTCAGTTCCGGGCGGAAAGCATCAACAAAGCATTCCAGAGCAACGGGTGTCCTTGGAC	3660

Db 3601 GGGCTCAGTTCCGGGGGAAAGATCACTAAAGCAATTCACAGACACGGGTCCCTCCGA 3660  
QY 3661 CAGCGCATCATATACCGGGCTCCATCAACCAAGGACGCGAGCTGCTCTGACAA 3720  
Db 3661 CAGCGCATCATATACCGGGCTCCATCAACCAAGGACGCGAGCTGAGTCTTACAA 3720  
QY 3721 GGGCAACATCACAGAGTATCGCGAGGACAGCCGAGTCTTGAACCGCGCGGGA 3780  
Db 3721 GGGCAACATCACAGAGTATCGCGAGGACAGCCGAGTCTTGAACCGCGCGGGA 3780  
QY 3781 GGAACAGCTGCCCAAGGGCCAGTCTTACGAAAGGCAAGAGGCGACGTCTTGCTTA 3840  
Db 3781 GGAACAGCTGCCCAAGGGCCAGTCTTACGAAAGGCAAGAGGCGACGTCTTGCTTA 3840  
QY 3841 TGAAGGTGATGTCTGTGATCCAGTGTCAAGAGGACGCGAGAGGCTCAAGACC 3900  
Db 3841 TGAAGGTGATGTCTGTGATCCAGTGTCAAGAGGACGCGAGAGGCTCAAGACC 3900  
QY 3901 CCCCCATGAGACGGCGCGCCCAAGCGCACTATGACATGATGAGAGGCGCGTGGCAG 3960  
Db 3901 CCCCCATGAGACGGCGCGCCCAAGCGCACTATGACATGATGAGAGGCGCGTGGCAG 3960  
QY 3961 AGCCATCTCTCAAGCAGATGAAAGTCTCATGGGCGGTGCCATTCGCGCGAGGACA 4020  
Db 3961 AGCCATCTCTCAAGCAGATGAAAGTCTCATGGGCGGTGCCATTCGCGCGAGGACA 4020  
QY 4021 CAGCCCCACACCTCAAGAGAGACACACATCCGCGGGTCAATCAACAAGGATCCC 4080  
Db 4021 CAGCCCCACACCTCAAGAGAGACACACATCCGCGGGTCAATCAACAAGGATCCC 4080  
QY 4081 TCGGTCTTACGTGAGGACACAGAGGACTACCTGCGTCGGAGGCGCAAGCTCTAAAGC 4140  
Db 4081 TCGGTCTTACGTGAGGACACAGAGGACTACCTGCGGTGGAGGCGCAAGCTCTAAAGC 4140  
QY 4141 GGAAGGACAGGCTTCGCGCCCAACCGCTCAAGGACCTGACCGAGGCTTAACAAGCGA 4200  
Db 4141 GGAAGGACAGGCTTCGCGCCCAACCGCTCAAGGACCTGACCGAGGCTTAACAAGCGA 4200  
QY 4201 GGGCCCTGGGCCCCCTGAAGCTAAGCGGCGCATGAGGGGCTGGTGAAGGA 4260  
Db 4201 GGGCCCTGGGCCCCCTGAAGCTAAGCGGCGCATGAGGGGCTGGTGAAGGA 4260  
QY 4261 GGGCGGCGCTTCATCATGAGATCCCGCGAGAGGCTGGGCAACGCCCGAGGTGCC 4320  
Db 4261 GGGCGGCGCTTCATCATGAGATCCCGCGAGAGGCTGGGCAACGCCCGAGGTGCC 4320  
QY 4321 CCTGGCCCCGCGGCTCTAAGAGGGCTTCATCAAGCAGGGCAACCGCTCAAGTACA 4380  
Db 4321 CCTGGCCCCGCGGCTCTAAGAGGGCTTCATCAAGCAGGGCAACCGCTCAAGTACA 4380  
QY 4381 CACCGGCGGCTCAACACTGGCTCCAAAAGACAGAGCTGCTCTGATGGGAGGCC 4440  
Db 4381 CACCGGCGGCTCAACACTGGCTCCAAAAGACAGAGCTGCTCTGATGGGAGGCC 4440  
QY 4441 CGGCGGACGTTCCCAACCGGTCAACCGCTGATGTATGCGCGACGCGCGGCACTGGA 4500  
Db 4441 CGGCGGACGTTCCCAACCGGTCAACCGCTGATGTATGCGCGACGCGCGGCACTGGA 4500  
QY 4501 AGGTGCTGTCTACGAGAGAGCTGAAGAGCGGCGACAGGACCGCGAGCTCGGGGG 4560  
Db 4501 AGGTGCTGTCTACGAGAGAGCTGAAGAGCGGCGACAGGACCGCGAGCTCGGGGG 4560  
QY 4561 CTCCATTTGGGGGGCGCGCGGTGATGTGCTGAGCTGGGTAAAGCGCGGCAAGGCC 4620  
Db 4561 CTCCATTTGGGGGGCGCGCGGTGATGTGCTGAGCTGGGTAAAGCGCGGCAAGGCC 4620  
QY 4621 CCTGACTTATGAGACACAGGGGCAACCTTTGCGGCGCACTTCCACAGAGTTCCCGGT 4680  
Db 4621 CCTGACTTATGAGACACAGGGGCAACCTTTGCGGCGCACTTCCACAGAGTTCCCGGT 4680  
QY 4681 GACCATGCGGAGGCCACGCGCGCTGAGAGAGGGCAGCTTTGTCAGCAAGGCATC 4740  
Db 4681 GACCATGCGGAGGCCACGCGCGCTGAGAGAGGGCAGCTTTGTCAGCAAGGCATC 4740

Db 4681 GACCATGCGGAGGCCACGCGCGCTGAGAGAGGGCAGCTTTGTCAGCAAGGCATC 4740  
QY 4741 CCAGGACGGAAAGCTGAGGTGACGCTCTGTGAGATCCCAAGTCCCGCACAGACCGT 4800  
Db 4741 CCAGGACGGAAAGCTGAGGTGACGCTCTGTGAGATCCCAAGTCCCGCACAGACCGT 4800  
QY 4801 GCCGAGACCAACCAACCCCATTCGCGCTATGAGCACTGTCTTGGGGCTGAGTGG 4860  
Db 4801 GCCGAGACCAACCAACCCCATTCGCGCTATGAGCACTGTCTTGGGGCTGAGTGG 4860  
QY 4861 CGTGAACCTGTATGSCAGACCATCCCTGGCTTGACCCCACTCATATCCCGGG 4920  
Db 4861 CGTGAACCTGTATGSCAGACCATCCCTGGCTTGACCCCACTCATATCCCGGG 4920  
QY 4921 CATCCCTGTGACGAGCGCTGCTACTATCACTGCGCCGACACTGGCCCCAACCCAC 4980  
Db 4921 CATCCCTGTGACGAGCGCTGCTACTATCACTGCGCCGACACTGGCCCCAACCCAC 4980  
QY 4981 CTACCGGACCTGTACCAACCTTACCTCATCGCGGCTACCCCGACAGCGCGCTGGA 5040  
Db 4981 CTACCGGACCTGTACCAACCTTACCTCATCGCGGCTACCCCGACAGCGCGCTGGA 5040  
QY 5041 GAAACGGAGACCATCATATGACTATGACTATCACTCGAGAGATGACCAACAAGGC 5100  
Db 5041 GAAACGGAGACCATCATATGACTATGACTATCACTCGAGAGATGACCAACAAGGC 5100  
QY 5101 CACGSCATGGCCGACGAGCTGATATGCTGAGGGGCTCTCGCCCGAGTCTCGCT 5160  
Db 5101 CACGSCATGGCCGACGAGCTGATATGCTGAGGGGCTCTCGCCCGAGTCTCGCT 5160  
QY 5161 GGCATCTCAACTACGCTGCGGCTCCCGAGGCAATCGACTGTGCCAAGTCCACACT 5220  
Db 5161 GGCATCTCAACTACGCTGCGGCTCCCGAGGCAATCGACTGTGCCAAGTCCACACT 5220  
QY 5221 GCTGTGCTGTGCCCCCGGACACAGAGGACCCAGGCAACCGCATGAGACCGCTTGCTTA 5280  
Db 5221 GCTGTGCTGTGCCCCCGGACACAGAGGACCCAGGCAACCGCATGAGACCGCTTGCTTA 5280  
QY 5281 CCTCCCAACCGGCCCAAGCTTCAAGCAGCGCCACAGAGTCCCATCTTCCCAAG 5340  
Db 5281 CCTCCCAACCGGCCCAAGCTTCAAGCAGCGCCCAAGAGTCCCATCTTCCCAAG 5340  
QY 5341 AGGTCAACACACTTGAACAAACCAACCAACGCTCTGTCCGAGCGGAGCGAGACG 5400  
Db 5341 AGGTCAACACACTTGAACAAACCAACCAACGCTCTGTCCGAGCGGAGCGAGACG 5400  
QY 5401 GATGAGAGAGGGGACCGGGAATCGGGAACGGGAAAGTCCATCTCAAGTCCACAGAC 5460  
Db 5401 GATGAGAGAGGGGACCGGGAATCGGGAACGGGAAAGTCCATCTCAAGTCCACAGAC 5460  
QY 5461 GGTGAGACGCAACCATCTGGAAGCTGTATCAGAGGAGAGCAGCGCAGCGGAG 5520  
Db 5461 GGTGAGAGCAGCAACCATCTGGAAGCTGTATCAGAGGAGAGCAGCGCAGCGGAG 5520  
QY 5521 CAGCGCGGGGTGGGGCAGACAGCGCGCTCCGCTTCCACTTCATGCCACAGCA 5580  
Db 5521 CAGCGCGGGGTGGGGCAGACAGCGCGCTCCGCTTCCACTTCATGCCACAGCA 5580  
QY 5581 CTGCGCCATCTTCCCTGGAACCCAGAGATCCCTCAAGAGAGACCCAGTGTTCACAA 5640  
Db 5581 CTGCGCCATCTTCCCTGGAACCCAGAGATCCCTCAAGAGAGACCCAGTGTTCACAA 5640  
QY 5641 CACAGGCAATGAAGGATATCATACCGCTGTGAAGCCAGAGCCACGATCTTGAAGTC 5700  
Db 5641 CACAGGCAATGAAGGATATCATACCGCTGTGAAGCCAGAGCCACGATCTTGAAGTC 5700  
QY 5701 CACCTCAACTTCTCAACCGTTGCGCAGCTGCCACATTTCCACCTGCACTGCC 5760  
Db 5701 CACCTCAACTTCTCAACCGTTGCGCAGCTGCCACATTTCCACCTGCACTGCC 5760  
QY 5761 ACTGGGCGGCAACCTTCATGAGGGGTACCTTAACCTTATGAGAGCCGTCTTGCCCA 5820  
Db 5761 ACTGGGCGGCAACCTTCATGAGGGGTCTACCTTAACCTTATGAGAGCCGTCTTGCCCA 5820



QY	5821	GGAGGCCCCCGGGGTCGCCCGGCGCAGAGCGGCCCGGAGCAGACACCGGCACTGCCTTCT	5880
DB	5821	GGAGGCCCCCGGGGTCGCCCGGCGCAGAGCGGCCCGGAGCAGACACCGGCACTGCCTTCT	5880
QY	5881	CGCCAGGCCCCCAGCGCGCTCCGGCTTGAGACCGCGCTCTCCCGACGAAGGGCTCGGA	5940
DB	5881	CGCCAGGCCCCCAGCGCGCTCCGGCTTGAGACCGCGCTCTCCCGACGAAGGGCTCGGA	5940
QY	5941	GCCCCGCGCCCTAAGGCTCTGTCCTTGAGCAGCGCACACTGCGCCGCAACCCCTGCGGA	6000
DB	5941	GCCCCGCGCCCTAAGGCTCTGTCCTTGAGCAGCGCACACTGCGCCGCAACCCCTGCGGA	6000
QY	6001	GAACTCGCACCTCAACGACCGCAGCCCGGACCCGCGGCGCACCTGCTCGGACTCGGA	6060
DB	6001	GAACTCGCACCTCAACGACCGCAGCCCGGACCCGCGGCGCACCTGCTCGGACTCGGA	6060
QY	6061	CCCCGACCGGGGAAAAAGACTCAAAAGTAAACCTTTTTCATCCAGGAACCTGGAATCTCCGTC	6120
DB	6061	CCCCGACCGGGGAAAAAGACTCAAAAGTAAACCTTTTTCATCCAGGAACCTGGAATCTCCGTC	6120
QY	6121	TCTGGGTTTACACAGCGCAGAGCTACAGCCCGGAGGGGTGAGGCCGCTCAGCCCTGAG	6180
DB	6121	TCTGGGTTTACACAGCGCAGAGCTACAGCCCGGAGGGGTGAGGCCGCTCAGCCCTGAG	6180
QY	6181	CTCAACCCAGTCTGACCCACGACAAAGGGGCTCCCAAGCACTTGGAAGACTCGACAG	6240
DB	6181	CTCAACCCAGTCTGACCCACGACAAAGGGGCTCCCAAGCACTTGGAAGACTCGACAG	6240
QY	6241	CCACCTGGAGGGGGAGCTGCGGCCAGCAGCCAGGCCCTGTGAAGCTTGGCGGGAGGC	6300
DB	6241	CCACCTGGAGGGGGAGCTGCGGCCAGCAGCCAGGCCCTGTGAAGCTTGGCGGGAGGC	6300
QY	6301	CGCCCACTCCCAACACCTGCGGGCGCTGCGCTGAGAGCCAGGCCCTCTCAACCCGCTGCT	6360
DB	6301	CGCCCACTCCCAACACCTGCGGGCGCTGCGCTGAGAGCCAGGCCCTCTCTCAACCCGCTGCT	6360
QY	6361	CCAGACCGCGCCAGAGGGTCAAAAGGTCAACACCGGGTGTCAACCTTGCGCCAGCATCAG	6420
DB	6361	CCAGACCGCGCCAGAGGGTCAAAAGGTCAACACCGGGTGTCAACCTTGCGCCAGCATCAG	6420
QY	6421	TGAGGTCATCAACACGACTACACCCCGGACCAACCAAGCAAGCTCAGCGCAACCCCTGTC	6480
DB	6421	TGAGGTCATCAACACGACTACACCCCGGACCAACCAAGCAAGCTCAGCGCAACCCCTGTC	6480
QY	6481	CGCCCCCTCTACTCTCTTCCCTTGGGGGACAGCTGCGCCGCTCTTGGAACCTTCGCGCCAC	6540
DB	6481	CGCCCCCTCTACTCTCTTCCCTTGGGGGACAGCTGCGCCGCTCTTGGAACCTTCGCGCCAC	6540
QY	6541	CAGTGACCTCTACTCTCCCGCGCCCGGACCAATGATGCCCCGCGCGTGGCTCCGCCACAG	6600
DB	6541	CAGTGACCTCTACTCTCCCGCGCCCGGACCAATGATGCCCCGCGCGTGGCTCCGCCACAG	6600
QY	6601	CGAAGGGGGCAAGAGTCTCCAGAGCCCAACAGACGTCTGGTCTTGGGTGTGTGAGGA	6660
DB	6601	CGAAGGGGGCAAGAGTCTCCAGAGCCCAACAGACGTCTGGTCTTGGGTGTGTGAGGA	6660
QY	6661	CGGATTTGAACTGTGTCTCCACCGGAGGGCATGACGAGCCAGAGGCACTCCCGGAATGTC	6720
DB	6661	CGGATTTGAACTGTGTCTCCACCGGAGGGCATGACGAGCCAGAGGCACTCCCGGAATGTC	6720
QY	6721	TGTGTACCCGCTGCTGTACCGGAAATGGGAAACAGACGGAAGCCAGAGAGTGGGCTCAA	6780
DB	6721	TGTGTACCCGCTGCTGTACCGGAAATGGGAAACAGACGGAAGCCAGAGAGTGGGCTCAA	6780
QY	6781	GTCCTCAGGCAACACAGCCAGCGCGCAAGCTTCTTTCAGCAAGCTGACCGAGACAACTC	6840
DB	6781	GTCCTCAGGCAACACAGCCAGCGCGCAAGCTTCTTTCAGCAAGCTGACCGAGACAACTC	6840
QY	6841	CGCCATGGTCAAGTCTCAAGAAAGCAAGATCAACAGAACTGAAACATCCCAACCGGAA	6900
DB	6841	CGCCATGGTCAAGTCTCAAGAAAGCAAGATCAACAGAACTGAAACATCCCAACCGGAA	6900

[illegible]

Db	7971	CCCTCCGCGCTCCCATCCGCTTACGCTCTGACAGATGAGCAGGCGCTGTCCAGCCC	8030
Qy	8041	CCAGNCGCGTGTCCGCGTCCCAAGACAGCTGCCCCAGCCCAAGAGATTGCTGAAACCA	8100
Db	8031	CCAGGCGCGTGTCCGCGTCCCAAGACAGCTGCCCCAGCCCAAGAGATTGCTGAAACCA	8090
Qy	8101	GTCAAGCCAGGTGGCGGACAAAGAGGCGGTGCGGCTGGGGGGAACGATGCTCCGA	8160
Db	8091	GTCAAGCCAGGTGGCGGACAAAGAGGCGGTGCGGCTGGGGGGAACGATGCTCCGA	8150
Qy	8161	GGACTGAGCTGTTTTTTTTCACACATGTTGCGGACGCGGTGGGAAGGAAGCAATGTA	8220
Db	8151	GGACTGAGCTGTTTTTTTTCACACATGTTGCGGACGCGGTGGGAAGGAAGCAATGTA	8210
Qy	8221	AATGATGATTGGTTTACAGGGTATATTTTATATCTTCAATGATTAATTAATCAGATGT	8280
Db	8211	AATGATGATTGGTTTACAGGGTATATTTTATATCTTCAATGATTAATTAATCAGATGT	8270
Qy	8281	TTACGCAAGAGAGACTTACCAAGATTAATCTGCTGTGCTTTTATCTGCTTACCG	8340
Db	8271	TTACGCAAGAGAGACTTACCAAGATTAATCTGCTGTGCTTTTATCTGCTTACCG	8330
Qy	8341	TTCAAGAGCGGTGTGCAAGGCGCAAGTGGTGAACCCCATCACTCCGAGAGCCAGGGGCG	8400
Db	8331	TTCAAGAGCGGTGTGCAAGGCGCAAGTGGTGAACCCCATCACTCCGAGAGCCAGGGGCG	8390
Qy	8401	GGGAGCTGCTGTCACGCGCCGCTGTGCTCCCTCCCTCCCTCCCTGCGGAGATGTA	8460
Db	8391	GGGAGCTGCTGTCACGCGCCGCTGTGCTCCCTCCCTCCCTCCCTGCGGAGATGTA	8450
Qy	8461	TTGATGCGGATTTCTGTGGCGCCCAATTTGGCGAGGGTGGTGTATTTCTGCTATTACCA	8520
Db	8451	TTGATGCGGATTTCTGTGGCGCCCAATTTGGCGAGGGTGGTGTATTTCTGCTATTACCA	8510
Qy	8521	CGTGTCTTAAATTAATAAAGCGAATTACTCCAAAAAATTAATAAATAAATAAATAA	8571
Db	8511	CGTGTCTTAAATTAATAAAGCGAATTACTCCAAAAAATTAATAAATAAATAAATAA	8561
RESULT 2	AF125672	8686 bp mRNA linear PRI 04-APR-1999	
LOCUS	AF125672	Homo sapiens silencing mediator of retinoic acid and thyroid hormone receptor extended isoform (SMRTE) mRNA, complete cds.	
DEFINITION	AF125672	homo receptor extended isoform (SMRTE) mRNA, complete cds.	
ACCESSION	AF125672.1	GI:4559297	
VERSION			
KEYWORDS			
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 8686)		
AUTHORS	Park, E.-J., Schreien, D.-J., Yang, M., Li, H., Li, L. and Chen, J.D.		
TITLE	SMRTE, a silencing mediator for retinoid and thyroid hormone receptors-extended isoform that is more related to the nuclear receptor corepressor		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3519-3524 (1999)		
REFERENCE	10097068		
PUBMED	2 (bases 1 to 8686)		
REFERENCE	Chen, J.D.		
AUTHORS	Direct Submission		
TITLE	Submitted (03-FEB-1999) Pharmacology and Molecular Toxicology, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655, USA		
JOURNAL			
FEATURES			
source	1..8686	Location/Qualifiers	
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
gene	1..8686		
	/gene="SMRTE"		

CDS	157..7680	/gene="SMRTE"		
		/function="transcriptional corepressor"		
		/note="SMRTE"		
		/codon_start=1		
		/product="silencing mediator of retinoic acid and thyroid hormone receptor extended isoform"		
		/protein_id="A022973.1"		
		/db_xref="GI:4559298"		
		/translation="MSGSTQDVQWRAETREPRPHSLSYEVQIARHTTDTVLLEYQHSRYASHLSRPSIIIPORRRPSILSRPSPDNESQELHLPESHSLYBELGKSMERIESRPRLELPDLRLRPSRLATGQPSGSDLTDRBSLTGKLEVPSPSPPHDPELELVPRLSKEBLIQMDRVDREITVNBQISKLKKQOQLSEBAKXPPREKPSPPIESKRRSLVOIITDENRKAARHITLGLDPQVELPYNPSPDRQHEHNTKIQAMRKKLILYFKRHHAKQWQKFCORYDDLAMEKEKVERIRINRRRKRKESKVRAYEKOPEIRKORLEJOEMQVQVGGSLSSPASEHEVEEITDGLSEQNLKQMLAVTPMLYADQDQRIKFTIMNGMLADPMKVYVLMQMMWMSHOKETPEKMQHKNFGLIASFLERKTVABCVLYIYITKKNENTKSLVRSYRRRGSQQQQQQQQQQQQQOPMPSIQEERDEKGTATANSQGRKGRITRSMANVANSBEAITPQQSALSMELNBSRRTBEMETAKGLIEGRNWSAIAWVSGSTVSOCKNFYNTKKQNLDEILQOHLKMEKSRNARKKKAPAAEBAAPPPVVEDEMEASGVTEEMEVEAEATVNSDYESIISPTEAAKADQNGPKRPATLIGADPPPGPPTPPEDIPAPTESTPASEATLATPPAPPPSAPPVYVKEKEKEETAAAPVYBEGEQKPPAAELAVDTGKAEBVKSCEBEAEGPSPAPKAGDAEAAEATAEALAEKKEGSGGATTAKSAGAPDSATGSCADYDEABGGDKNRLSPRSLITPTGDPANASPOKELDKQKORAAIPLIQTVHVRPREDAAPTKAPAPPPONLQEPESDAPQOQPSRCKSSPAPADKEAKVPFPAPAAEQKLPQDPFMTGSLPVPVPRVITKASPHAPPSAFVPPGHPLEGLADTARVLPKLPISNPPRLISSAKHPSVLEROIATISQGSVOLHVSASHAKAPVPTMGALPMDPKLAPPSGVKQDOLSPRQAGPPRESIGVPTAQAASLTRTLAGSVSGSITKQIPPTRVSDSAITTRGSLTHGTPADVILKGTITLITGDSRLDRGSDSLPKGHVILEGKGHVLSYEGMSVYQCKEDGRSSGPPHATAKRYVDMWEGVSAISASIEGLKMGRAIPREHSPHNLKEQHNRIGSLTQCIPSYVGAQDYLREKALKRGTAPPPRSLTEAVKTPQALPLKLPANHELVAIVEAKESIHIREELRHLREPLARPLKEGSITQGTPLKDTGASTTSGKDVSLIGSPERTPPVPHLPDVMDADALBACYEESLSKRPCTASSSGGSIARCAPVYVPELCKRPSPLTYRHHGAPRANHLRGSVYREPTRLQSGSLSSKASQDRKLTSTPELAKSPSTVPEHNPISYEHLKRVSGVDLVRSHPLAPDPSIIPRGILPDAAAVYVPHLAPPEVPHLYPLIGYPTAALENRQTIINDYITSGQHNTATMAAQADMLRGISPRESSIALNVAAPGPIIDLSQVPLPLVLPPTGPPTATMDRLAYLPTAPQPSHSSSSPLSGGPTLPTPTTSSSEBRDRERDRERBKSILITSTTVHAPIMRGTOSSGSSGSGGSSSRPASHSHTSHAHGNSPISTRQDALQORPSVHTMCKKGIITVPEKPTVLTSTSSPVRAAITPPRATCPVIGTUDVYPTLMEVPLPKAPARARERPRATGHAFLAKPPRSGLEPPASPSKSGEPRLVPPVSGHATATPAKMLPHNASDPAPPPASADPHKEXTQSKPESIQELRLSLGYHSSYSEGVVPPVSPSLTHKGLKXHLIEKLSHLEGELRPKQPEVLGGAALPLPHLRPLEBQSPSSPLQTPAGVKGHQTVAQLAHSIVITQDVTNRHNPQOLSAPLAPLPPGASCPGLTLDRPSDLYLPPDGAAPRGSFHS EGGKRSRPNNTSVLGGEDIEPVSPEPGMTERCHSRSAVYPLLYRGGEDTERMGSKSPNTSQQPAPFSKLTESNANVSKKQIKINKKLTNNRNPETINSQETELFMNPAITGTGMYTASQAVQDHASTMGLXIIKALMGKXDQWEEPSLANAFNPINASASLPAAMPITTAADGSDHTLTPSGGGAFAKVGSPSSSRKAKSPAGLSADRPPVSSVHSQDCNRRTPLNVRWEDRPSGASIPFPYPLIMLQGVMSPPPGLPAGSGPLAGPHNADDEPKPLCSQYETLSDSE"		
Query Match	97.2%	Score 8324.6;	DB 8; Length 8686;	
Best local Similarity	98.6%	Pred. No. 0;		
Matches 8472;	Conservative 1;	Mismatches 34;	Indels 88; Gaps 4;	
Qy	1	CATGTGGGCTTCACACAGCTTTGTGGCAACAGCTGAGGGCCACTGAGCCCGCTTACC	60	
Db	156	CATGTGGGCTTCACACAGCTTTGTGGCAACAGCTGAGGGCCACTGAGCCCGCTTACC	215	
Qy	61	GCCCCAGGCTTTCTTCCAGTGCAGATGCGCCGAGCGACACAGGAGTGGGCTCT	120	
Db	216	GCCCCAGGCTTTCTTCCAGTGCAGATGCGCCGAGCGACACAGGAGTGGGCTCT	275	
Qy	121	GGAGTACAGACCACTCCGCGACTATGCTCCACCTGTGCGCGGCTCCATATCCA	180	
Db	276	GGAGTACAGACCACTCCGCGACTATGCTCCACCTGTGCGCGGCTCCATATCCA	335	
Qy	181	GCCCCAGGCGGAGGCGCTTCTGCTGTCTGAGTTCCAGCCCGGAATGAACGTTCCA	240	

## ORIGIN

Query Match 97.2%; Score 8324.6; DB 8; Length 8686;  
 Best Local Similarity 98.6%; Pred. No. 0;  
 Matches 8472; Conservative 1; Mismatches 34; Indels 88; Gaps 4;

Db 336 GCCCAGCGCGGAGGCCCTCCCTGCTGTCTGAGTTCCAGCCCGGAGATGACGATCCCA 395  
QY 241 GGAGCTCAACCTGCGCGCAGAGTCCCACTATACCTGCGCGAGCTGGGAGATCAAGAT 300  
Db 396 GGAGCTCAACCTGCGCGCAGAGTCCCACTATACCTGCGCGAGCTGGGAGATCAAGAT 455  
QY 301 GGAGTTCAATTGAAAGCAGCGCCTCGCTGAGCTGCTGACCTGACCTCCCTGCTGAC 360  
Db 456 GGAGTTCAATTGAAAGCAGCGCCTCGCTGAGCTGCTGACCTGACCTCCCTGCTGAC 515  
QY 361 GTCAACCTGCTGCTGCGCAACCGGCGAGCTGCGGATCTGAAAGCTTCAACAGACCTGAG 420  
Db 516 GTCAACCTGCTGCTGCGCAACCGGCGAGCTGCGGATCTGAAAGCTTCAACAGACCTGAG 575  
QY 421 CCTGACGGGCAAGCTGGAACCGGCTGTCTCCCGCAGCGCCCGCACTGACCTTGAGCT 480  
Db 576 CCTGACGGGCAAGCTGGAACCGGCTGTCTCCCGCAGCGCCCGCACTGACCTTGAGCT 635  
QY 481 GGAGCTGTGCGCGCAACCGCTGTCTCAAGAGAGAGCTGATCCAGAACATGAGACCGCTGGA 540  
Db 636 GGAGCTGTGCGCGCAACCGCTGTCTCAAGAGAGAGCTGATCCAGAACATGAGACCGCTGGA 695  
QY 541 CGAGAGATCAACATGCTGAGAGCAAGATCTTTAAGCTGAGAGAGAGAGCAAGCACT 600  
Db 696 CGAGAGATCAACATGCTGAGAGCAAGATCTTTAAGCTGAGAGAGAGAGCAAGCAAGCT 755  
QY 601 GGAGAGAGAGCTGCGCAACCGCGCGAGCTGAGAACCGCGCTGCAACCGCGCGCAATCGA 660  
Db 756 GGAGAGAGAGCTGCGCAACCGCGCGAGCTGAGAACCGCGCTGCAACCGCGCGCAATCGA 815  
QY 661 GTCGAAGCACCAGAGCTGCTGAGATCACTACGACGAGAACCGGAGAGAGCTGAGAG 720  
Db 816 GTCGAAGCACCAGAGCTGCTGAGATCACTACGACGAGAACCGGAGAGAGCTGAGAG 875  
QY 721 TGCACATCGAATTTCTGAAAGCGCTGAGGCGCCAGCTGAGAGCTGCGCTGTACAAACG 780  
Db 876 TGCACATCGAATTTCTGAAAGCGCTGAGGCGCCAGCTGAGAGCTGCGCTGTACAAACG 935  
QY 781 CTCGAGCACCAGAGCTGATATGAGAACATCAAAATTAACAGAGCATGCGAGAGAGCT 840  
Db 936 CTCGAGCACCAGAGCTGATATGAGAACATCAAAATTAACAGAGCATGCGAGAGAGCT 995  
QY 841 AATCTTGATCTTCAAGAGAGAGATCAAGCTCGGAAACAATGAGACAGAGATTTCTGCA 900  
Db 996 AATCTTGATCTTCAAGAGAGAGATCAAGCTCGGAAACAATGAGAGCAAGATTTCTGCA 1055  
QY 901 GCGCTATGACCAAGCTCATGAGGCTTGGAGAAAGAGTGAAGCGCATCGAGAACAAACC 960  
Db 1056 GCGCTATGACCAAGCTCATGAGGCTTGGAGAAAGAGTGAAGCGCATCGAGAACAAACC 1115  
QY 961 GCGCGCGCGCGCCAGAGAGAGAGTGGCGGATCTACGAAACAAGTTCCTGAGAT 1020  
Db 1116 GCGCGCGCGCGCCAGAGAGAGAGTGGCGGATCTACGAGAACAGTTCCTGAGAT 1175  
QY 1021 CCGCAAGCAGCGCAGAGCTGAGAGCGCATGACAGAGAGAGTGGCGCAGCGGCGAGTGG 1080  
Db 1176 CCGCAAGCAGCGCAGAGCTGAGAGCGCATGACAGAGAGAGTGGCGCAGCGGCGAGTGG 1232  
QY 1081 GCTGTTCATGTGCGCGCGCCGACGAGAGCAAGAGTGTCAAGATCATGTGAGCTTCTC 1140  
Db 1233 GCTGTTCATGTGCGCGCGCCGACGAGAGCAAGAGTGTCAAGATCATGTGAGCTTCTC 1292  
QY 1141 AAGAGAGAGAGAGAGTGGAGAGAGATGGCGCAGCTGCGCGCTGATCCCGCCATGCTTGA 1200  
Db 1293 AAGAGAGAGAGAGAGTGGAGAGAGATGGCGCAGCTGCGCGCTGATCCCGCCATGCTTGA 1352  
QY 1201 CGAGCTGACAGCAGCGCATCAAGTTCAATCAACATGAAACGAGCTTATGGCCGACCCAT 1260  
Db 1353 CGAGCTGACAGCAGCGCATCAAGTTCAATCAACATGAAACGAGCTTATGGCCGACCCAT 1412  
QY 1261 GAAGGTGTACAAAGACCGCAGGTCATGAAATGTGAGTGAAGCAGAGAGAGAGACTTT 1320  
Db 1413 GAAGGTGTACAAAGACCGCAGGTCATGAAATGTGAGTGAAGCAGAGAGAGAGACTTT 1472

QY 1321 CCGGAGAAAGTTATGACAGATCCCAAGAACTTTGGCCTGATGATCATTTCTGAGAG 1380  
Db 1473 CCGGAGAAAGTTATGACAGATCCCAAGAACTTTGGCCTGATGATCATTTCTGAGAG 1532  
QY 1381 GAAGCAGTGGCTGAGTGGCTCTCTTATTAATTAATTAATTAAGAGATGAGAACTATTA 1440  
Db 1533 GAAGCAGTGGCTGAGTGGCTCTCTTATTAATTAATTAATTAAGAGATGAGAACTATTA 1592  
QY 1441 GAGCTGTGAGAGCGAGAGCTATGAGGCGCGCGCAAGAGCAGCAGCAAGAGAGAGCA 1500  
Db 1593 GAGCTGTGAGAGCGAGAGCTATGAGGCGCGCGCAAGAGCAGCAGCAAGAGAGAGCA 1552  
QY 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
Db 1653 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1712  
QY 1561 AGATGAG 1620  
Db 1713 AGATGAG 1772  
QY 1621 CGACAG 1680  
Db 1773 CGACAG 1832  
QY 1681 GAAGAGAGCTGAGCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740  
Db 1833 GAAGAGAGCTGAGCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1892  
QY 1741 CCGCATCAACCGCTCAATGAGCTTAATGAGGCAACAGCAGAGAGAGAGAGAGAGAG 1800  
Db 1893 CCGCATCAACCGCTCAATGAGCTTAATGAGGCAACAGCAGAGAGAGAGAGAGAGAG 1952  
QY 1801 GAGCGCGAGCTGAGCTCCATGAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860  
Db 1953 GAGCGCGAGCTGAGCTCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2012  
QY 1861 GAG 1920  
Db 2013 GAG 2072  
QY 1921 GGTGGGCTCAAGAGCTGTGTGCGAGTGTAGAACTTTCTAATTAAGAGAGAGAGAG 1980  
Db 2073 GGTGGGCTCAAGAGCTGTGTGCGAGTGTAGAACTTTCTAATTAAGAGAGAGAGAG 2132  
QY 1981 GAACCTGATGAGATTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040  
Db 2133 GAACCTGATGAGATTTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2192  
QY 2041 GAG 2100  
Db 2193 GAG 2252  
QY 2101 GAGAGATGAGAGATGAG 2160  
Db 2253 GAGAGATGAGAGATGAG 2212  
QY 2161 GGTGTAAGCTTATCATGCTCTTGGAGATGAGGTGCCAGAGAGAGAGATGAGTGGCCAG 2220  
Db 2313 GGTGTAAGCTTATCATGCTCTTGGAGATGAGGTGCCAGAGAGAGAGATGAGTGGCCAG 2320  
QY 2221 CACTGTCAACACAGCTCAG 2280  
Db 2322 CACTGTCAACACAGCTCAG 2381  
QY 2281 GAGCAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340  
Db 2382 GAGCAGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2441  
QY 2341 AGGCGCAG 2400  
Db 2442 AGGCGCAG 2501

[illegible]

Db	3582	GCCCCCTGCCCCATGAGACCCGAAAAAGCTGGCACCCCTTACGCGGAGTGAAGCAGAGCAGCT	3641
Qy	3517	GTCCCAACGAGAGCCAGAGCTGGGACACCGGAGAGCTGAGAGGTGCCCAAGCTCCAGAGAGC	3576
Db	3642	GTCCCCACGGGAGCCAGAGCTTGGGACACCGGAGAGCTGGGGGTGGCCACAGGCCAGGAGGC	3701
Qy	3577	GTCCGTGCTGACAGAGGACAGCTCTGGGGCTCAGTTCCGGGGCGAAGACATCACAAAAGCAT	3636
Db	3702	GTCCGTGCTGAGAGGAGCAGCTCTGGGGCTCAATTTCCGGGCGAAGACATCACTAAAGGAT	3761
Qy	3637	TCCACAGACACGAGAGTCCCTCGGACAGGCGCAATCAATACCGGGCTCATCACCCACG	3696
Db	3762	TCCACAGACACGAGAGTCCCTCGGACAGGCGCATCAATACCGGGCTCATCACCCACG	3821
Qy	3697	CACGCAAGCTGACGTCTGTACAAAGGACCATCAACAAGATCATCGGAGAGCAGGCC	3756
Db	3822	CACGCAAGCTGACGTCTGTACAAAGGACCATCAACAAGATCATCGGAGAGCAGGCC	3881
Qy	3757	GAGTCGCTTGAACCGGAGCGGGGAGGACAGCTTGGCCAAAGGGCCAGCTCATCTACGAAG	3816
Db	3882	GAGTCGCTTGAACCGGAGCGGGGAGGACAGCTTGGCCAAAGGGCCAGCTCATCTACGAAG	3941
Qy	3817	CAAGAAAGGACAGCTTGTGCTTACGAGGAGTGAATCTGTAACCAAGTCTCCAAAGA	3876
Db	3942	CAAGAAAGGACAGCTTGTGCTTACGAGGAGTGAATCTGTAACCAAGTCTCCAAAGA	4001
Qy	3877	GACGCGCAGAGCAGCTTCAGACCCCCCATGAGACGGCCGCCCCAAGCGCACCTATGA	3936
Db	4002	GACGCGCAGAGCAGCTTCAGACCCCCCATGAGACGGCCGCCCCAAGCGCACCTATGA	4061
Qy	3937	CATGATGAGAGGCGCGGTGGGACAGACCATCTCTCAGCAGCATCCAAAGTCTCATGGG	3996
Db	4062	CATGATGAGAGGCGCGGTGGGACAGACCATCTCTCAGCAGCATCCAAAGTCTCATGGG	4121
Qy	3997	CGGTGCATATCCGCGCGGAGCGACAGAGCCCCACCAACCTCAAAGGCGACCAATCCG	4056
Db	4122	CGGTGCATATCCGCGCGGAGCGACAGAGCCCCACCAACCTCAAAGGCGACCAATCCG	4181
Qy	4057	CGGGTCATATCACAAAGGAGTCCCTCGGTCTTACGTGAGGACAGAGAGCTACCTTCCG	4116
Db	4182	CGGGTCATATCACAAAGGAGTCCCTCGGTCTTACGTGAGGAGACAGAGAGCTACCTTCCG	4241
Qy	4117	TGCGAGAGGCCAGTCTTAAAGCGGAGGAGCACGCTCTCCGCCCTTACCGGGA	4176
Db	4242	TGCGAGAGGCCAGTCTTAAAGCGGAGGAGCACGCTCTCCGCCCTTACCGGGA	4301
Qy	4177	CTGACCCGAGGCTTAAAGACGACGAGCCCTTGGGCCCCCTGAAGCTGAAGCGGGCCATGA	4236
Db	4302	CTGACCCGAGGCTTAAAGACGACGAGCCCTTGGGCCCCCTGAAGCTGAAGCGGGCCATGA	4361
Qy	4237	GGGCTGTGTGACCAAGGTGAAGAGGCGGGCGGTCTCATTCATGAGATCCCGCGAGGA	4296
Db	4362	GGGCTGTGTGACCAAGGTGAAGAGGCGGGCGGTCTCATTCATGAGATCCCGCGAGGA	4421
Qy	4297	GCTGCGGACACGCGCCGAGCTGCCCCCTGGCCCCCGGAGCTGCTCAAGAGGCTTCATCAC	4356
Db	4422	GCTGCGGACACGCGCCGAGCTGCCCCCTGGCCCCCGGAGCTGCTCAAGAGGCTTCATCAC	4481
Qy	4357	GCAAGGACACCCGCTTCAATGACACCGGCGGCTTCACCATGCTTCCAAAAGCAGGA	4416
Db	4482	GCAAGGACACCCGCTTCAATGACACCGGCGGCTTCACCATGCTTCCAAAAGCAGGA	4541
Qy	4417	CGTACGCTCCCTATCGAGGAGCCCCGCGAGAGCTTCCACACCGGTGCACCCGCTGATGT	4476
Db	4542	CGTACGCTCCCTATCGAGGAGCCCCGCGAGAGCTTCCACACCGGTGCACCCGCTGATGT	4601
Qy	4477	GATGCGCGACGCCCCGCGACCTGGAACGTGCTTGTACGAGAGAGAGCTTGAAGCGCGCC	4536
Db	4602	GATGCGCGACGCCCCGCGACCTGGAACGTGCTTGTACGAGAGAGAGCTTGAAGAGCGCGCC	4661
Qy	4537	AGGAGACCGCCAGCAGCTTCGAGGAGCTCCATTGCGCGCGAGCCCGGCTCATTTGTGCTTGA	4596

Db 4662 AGGAGCCGACGAGCTCGGGGGGCTCCATTGCGCGGCGCCCGATCATTTGCTGTA 4721  
QY 4597 GCTGGGTAAGCCGCGGAGAGCCCTGACCTTAGAGACACGAGGACACCTTTGGCGG 4656  
Db 4722 GCTGGGTAAGCCGCGGAGAGCCCTGACCTTAGAGACACGAGGACACCTTTGGCGG 4781  
QY 4657 CCACTTCCACGAGGTTGCGCCGCTGACATGACGAGAGCCACGCGGCTTGAGAGAGG 4716  
Db 4782 CCACTTCCACGAGGTTGCGCCGCTGACATGAGGAGAGCCACGCGGCTTGAGAGAGG 4841  
QY 4717 CAGCTTTTGTCCAGAGGATCCCAAGACCGAAAGCTGACTCGACGCTGTGAGAT 4776  
Db 4842 CAGCTTTTGTCCAGAGGATCCCAAGACCGAAAGCTGACTCGACGCTGTGAGAT 4901  
QY 4777 CGCAGAGTCCCGCAGAGACCGTGCCTGAGCAACCCACACCCATCGGCTTAGTA 4836  
Db 4902 CGCAGAGTCCCGCAGAGACCGTGCCTGAGCAACCCACACCCATCGGCTTAGTA 4961  
QY 4837 GCACTGTGCGGGGCGTAGAGTGGCGTGAACCTGTATCGAGCGACATCCCGCTGACTT 4896  
Db 4962 GCACTGTGCGGGGCGTAGAGTGGCGTGAACCTGTATCGAGCGACATCCCGCTGACTT 5021  
QY 4897 CGACCCCACTCTCATACCCCGCGCATTCCTCTGAGACGCGGCTGCTTAATCTGCC 4956  
Db 5022 CGACCCCACTCTCATACCCCGCGCATTCCTCTGAGACGCGGCTGCTTAATCTGCC 5081  
QY 4957 CGAGACCTGGGCGCCCAACCCCACTAACCGGCACTGTACCCACTCACTCCGCGG 5016  
Db 5082 CGAGACCTGGGCGCCCAACCCCACTAACCGGCACTGTACCCACTCACTCCGCGG 5141  
QY 5017 CTACCCCGACACGCGCGCGCTGAGAAACGCGCAGACCATCATATGACTACATCACTC 5076  
Db 5142 CTACCCCGACACGCGCGCGCTGAGAAACGCGCAGACCATCATATGACTACATCACTC 5201  
QY 5077 GCAAGCATGACCAACAACACGCGCACCGCATAGCCCAAGAGTGTATGCTGAGAGG 5136  
Db 5202 GCAAGCATGACCAACAACACGCGCACCGCATAGCCCAAGAGTGTATGCTGAGAGG 5261  
QY 5137 CCTCTGCGCGCGGAGTCCCTGCTGACATCACTACGCTGCGGGTCCCGAGGACATCAT 5196  
Db 5262 CCTCTGCGCGCGGAGTCCCTGCTGACATCACTACGCTGCGGGTCCCGAGGACATCAT 5321  
QY 5197 CGACCTGTCCCAAGTCCACACCTGCTGTGCTGTGCCCCGACACACAGGCAACCCACG 5256  
Db 5322 CGACCTGTCCCAAGTCCACACCTGCTGTGCTGTGCCCCGACACACAGGCAACCCACG 5381  
QY 5257 CACCGCATGAGACCGCTTGTCTTACTCTCCACCGCGGCCAGCCCTTGACAGCCGCCA 5316  
Db 5382 CACCGCATGAGACCGCTTGTCTTACTCTCCACCGCGGCCAGCCCTTGACAGCCGCCA 5441  
QY 5317 CAGAGCTCCCACTTCCCGAGAGGTCCAAACAATTGACAAACCAACCAACCAACGTC 5376  
Db 5442 CAGAGCTCCCACTTCCCGAGAGGTCCAAACAATTGACAAACCAACCAACCAACGTC 5501  
QY 5377 CTCTGTCCAGCGGAGACGAGACCGGATCGAGACGCGGACCGGGATCGGGAGCGGGAAAA 5436  
Db 5502 CTCTGTCCAGCGGAGACGAGACCGGATCGAGACGCGGACCGGGATCGGGAGCGGGAAAA 5561  
QY 5437 GTTCATCTTCAAGTCCACACGAGGTGAGACGACCATCTGAGAACCTTGTAACA 5496  
Db 5562 GTTCATCTTCAAGTCCACACGAGGTGAGACGACCATCTGAGAACCTTGTAACA 5621  
QY 5497 GGAAGCAACCGGAGAGGCGAGACGCGCGGGGTTGGGGGAGAGAGACCGCGCCCG 5556  
Db 5622 GGAAGCAACCGGAGAGGCGAGACGCGCGGGGTTGGGGGAGAGAGACCGCGCCCG 5681  
QY 5557 CTTCACCTCCATGACCACACGACTCGCCATCTCCCTCGAGACCAAGATGCTCTCA 5616  
Db 5682 CTTCACCTCCATGACCACACGACTCGCCATCTCCCTCGAGACCAAGATGCTCTCA 5741  
QY 5617 GCAAGACCAAGTGTCTTCAACAACAGGATGAAAGGTATCATCACTGCTGAGAGCC 5676  
Db 5742 GCAAGACCAAGTGTCTTCAACAACAGGATGAAAGGTATCATCACTGCTGAGAGCC 5801

QY 5677 CAGCAAGCCCAAGGATCTGAGGTCAACTTCACACTCTCTCAACCCGTTGCCCCAGCTGCAC 5736  
Db 5802 CAGCAAGCCCAAGGATCTGAGGTCAACTTCACACTCTCTCAACCCGTTGCCCCAGCTGCAC 5861  
QY 5737 ATTCCCACTGCAACCACTGCACTGGCGGACCTCTGATGAGGCTTAACCTTACCT 5796  
Db 5862 ATTCCCACTGCAACCACTGCACTGGCGGACCTCTGATGAGGCTTAACCTTACCT 5921  
QY 5797 CATGAGCCCGCTTTCCTGCGCAAGAGGCCCCCGGGTCCCGGACAGAGGCGCCG 5856  
Db 5922 CATGAGCCCGCTTTCCTGCGCAAGAGGCCCCCGGGTCCCGGACAGAGGCGCCG 5981  
QY 5857 AGCAGACACCGGACATGACTCTCTCGCCAAAGCCCGACCGGCTCCGGGCTGAGCGCG 5916  
Db 5982 AGCAGACACCGGACATGACTCTCTCGCCAAAGCCCGACCGGCTCCGGGCTGAGCGCG 6041  
QY 5917 CTCCTCCCCAGCAAGGCTCGAGGCTCCGAGCCCGGCTTAAGTGCCTCTGTCTTGGCAAGC 5976  
Db 6042 CTCCTCCCCAGCAAGGCTCGAGGCTCCGAGCCCGGCTTAAGTGCCTCTGTCTTGGCAAGC 6101  
QY 5977 CACCATGCGCGGACCCCTGCGAAGAACTTCGCACTCAACGCAAGCCCGGACCGCGC 6036  
Db 6102 CACCATGCGCGGACCCCTGCGAAGAACTTCGCACTCAACGCAAGCCCGGACCGCGC 6161  
QY 6037 GCGGCACTGCTCGGCTCGGACCCGACACCGGAAAAAGACTCAAGTAAACCTTTTC 6096  
Db 6162 GCGGCACTGCTCGGCTCGGACCCGACACCGGAAAAAGACTCAAGTAAACCTTTTC 6221  
QY 6097 CATCAGAAACTGGAATCTCGTTCTGTGGTTTACAACGCGAGAGTTACAGCCCGAAGG 6156  
Db 6222 CATCAGAAACTGGAATCTCGTTCTGTGGTTTACAACGCGAGAGTTACAGCCCGAAGG 6281  
QY 6157 GGTGAGCCCGTCAAGCCCTGTGAGTCAACCACTCTGACCCCAAGACAAAGGAGGTTCCCA 6216  
Db 6282 GGTGAGCCCGTCAAGCCCTGTGAGTCAACCACTCTGACCCCAAGACAAAGGAGGTTCCCA 6341  
QY 6217 GCACTGGAAGAGCTGACAAAGAGCACTGAGGAGGAGCTGCGGCCCAAGACGACAG 6276  
Db 6342 GCACTGGAAGAGCTGACAAAGAGCACTGAGGAGGAGCTGCGGCCCAAGAGACGAG 6401  
QY 6277 CCCCCTGAAGCTTGGCGGAGAGCGGCCCACTTCCCAACTTGCAGCGCTGCTGAGAG 6336  
Db 6402 CCCCCTGAAGCTTGGCGGAGAGCGGCCCACTTCCCAACTTGCAGCGCTGCTGAGAG 6461  
QY 6337 CCAAGCCCTGTCAAGCCCGCTGCTCAAGACCGGCCCAAGGAGTCAACAGCGGCT 6396  
Db 6462 CCAAGCCCTGTCAAGCCCGCTGCTCAAGACCGGCCCAAGGAGTCAACAGCGGCT 6521  
QY 6397 GGTCAACCTGGCCAGACATGATGAGGTCAACACAGAACTAACCCCGGACCAACC 6456  
Db 6522 GGTCAACCTGGCCAGACATGATGAGGTCAACACAGAACTAACCCCGGACCAACC 6581  
QY 6457 ACAGCAGCTCAAGGCAACCCCTGCGCGCCCTCTTACTCTTCTGAGGCGAGCTGCC 6516  
Db 6582 ACAGCAGCTCAAGGCAACCCCTGCGCGCCCTCTTACTCTTCTGAGGCGAGCTGCC 6641  
QY 6517 CTTCTTGAACCTTCGCGCGCCCAACGTAACCTTAACCTTCGCGCCCGGACATGCTGC 6576  
Db 6642 CTTCTTGAACCTTCGCGCGCCCAACGTAACCTTAACCTTCGCGCCCGGACATGCTGC 6701  
QY 6577 CCGGCGCGGTGCTCCCGCAACGAGAAAGGAGGAGGTCTCCAGAGCAAAAGAGC 6636  
Db 6702 CCGGCGCGGTGCTCCCGCAACGAGAAAGGAGGAGGTCTCCAGAGCAAAAGAGC 6761  
QY 6637 GTTCAGTCTTGGGTGTGTGAGACGTTAATTGAACCTGTGTCCCAACGAGAGGCAATGAC 6696  
Db 6762 GTTCAGTCTTGGGTGTGTGAGACGTTAATTGAACCTGTGTCCCAACGAGAGGCAATGAC 6821  
QY 6697 GGAAGCAAGGCACTCCCGAGAGTGTGTATCCCGCTGTGTACCCGGAATGGGAAACAGAC 6756  
Db 6822 GGAAGCAAGGCACTCCCGAGAGTGTGTATCCCGCTGTGTACCCGGAATGGGAAACAGAC 6881

QY 6757 GGAGCCGAGAGATGGGCTCCAGTCTTCAGAGCAACACAGCCAGCCGACCTTCTT 6816  
| | | | |  
DB 6882 GGAGCCGAGAGATGGGCTCCAGTCTTCAGAGCAACACAGCCAGCCGACCTTCTT 6941  
| | | | |  
QY 6817 CAGCAAGCTGACCGAGAGCAATCCGCGCATGGTCAAGTCCAGAGCAAGATCAACA 6876  
| | | | |  
DB 6942 CAGCAAGCTGACCGAGAGCAATCCGCGCATGGTCAAGTCCAGAGCAAGATCAACA 7001  
| | | | |  
QY 6877 GAAGCTGAACACCCCAACCCGGAATGAGCCTGAATATCAATATCAGCAGCTGGACGGA 6936  
| | | | |  
DB 7002 GAAGCTGAACACCCCAACCCGGAATGAGCCTGAATATCAATATCAGCAGCTGGACGGA 7061  
| | | | |  
QY 6937 GATCTTCAATATGCGCGCATATCAGCGGAACAG3CCTTATGACTTATGAAAGCCAGCGGT 6996  
| | | | |  
DB 7062 GATCTTCAATATGCGCGCATATCAGCGGAACAG3CCTTATGACTTATGAAAGCCAGCGGT 7121  
| | | | |  
QY 6997 GAGAGAACTGCGACACCAACATGGGGGTGAGAGCCATATTTAAGAAAGCACTATGGG 7056  
| | | | |  
DB 7122 GAGAGAACTGCGACACCAACATGGGGGTGAGAGCCATATTTAAGAAAGCACTATGGG 7181  
| | | | |  
QY 7057 TAAATATGACCAAGTGGGAAGATCCCGCGCTCAGCGCCAAATGCTTTTAAACCTCTGA 7116  
| | | | |  
DB 7182 TAAATATGACCAAGTGGGAAGATCCCGCGCTCAGCGCCAAATGCTTTTAAACCTCTGA 7241  
| | | | |  
QY 7117 TGCAGATGCGAGCTTGCGCTGTATGCGCATTAACCGCTGCTGACGAGAGTGAACA 7176  
| | | | |  
DB 7242 TGCAGATGCGAGCTTGCGCTGTATGCGCATTAACCGCTGCTGACGAGAGTGAACA 7301  
| | | | |  
QY 7177 CACACTCAGCTGCGCAGAGTGGGGGGGGAAGGCCAAAGTCTGTGACAGACCCAGACCGG 7236  
| | | | |  
DB 7302 CACACTCAGCTGCGCAGAGTGGGGGGGGAAGGCCAAAGTCTGTGAGAGACCCAGACCGG 7361  
| | | | |  
QY 7237 AAAAGCCAAGTCCCGGCGCCCGGAGCTGAGCATCTGGGGAACCGGACACCTCTGTCTCTC 7296  
| | | | |  
DB 7362 AAAAGCCAAGTCCCGGCGCCCGGAGCTGAGCATCTGGGGAACCGGACACCTCTGTCTCTC 7421  
| | | | |  
QY 7297 AGTGCATCTGGAGGAGACTGCACACCGCGGACCGCGCTCACCAACCGGTGTGGAGGA 7356  
| | | | |  
DB 7422 AGTGCATCTGGAGGAGACTGCACACCGCGGACCGCGCTCACCAACCGGTGTGGAGGA 7481  
| | | | |  
QY 7357 CAGGCGCTGCTCGAGAGTTCACACGCAATCCCTCAACACCCCTGATATAGTGGCTGGA 7416  
| | | | |  
DB 7482 CAGGCGCTGCTCGAGAGTTCACACGCAATCCCTCAACACCCCTGATATAGTGGCTGGA 7541  
| | | | |  
QY 7417 GCGCGGATGATGAGCTTCCCAACCCCAACCGGAGCTCCCGCGGAGCGGAGCCCTCTCCG 7476  
| | | | |  
DB 7542 GCGCGGATGATGAGCTTCCCAACCCCAACCGGAGCTCCCGCGGAGCGGAGCCCTCTCCG 7601  
| | | | |  
QY 7477 TGGCCCCCAACACGCTGGAGACGAGAGCCCAAGCCACTGTCTGTGCAATGAGAC 7536  
| | | | |  
DB 7602 TGGCCCCCAACACGCTGGAGACGAGAGCCCAAGCCACTGTCTGTGCAATGAGAC 7661  
| | | | |  
QY 7537 ACTCTCCAACAGCGAGTGACTGAGAACAGGGGCGGGGGGGGGGGCGGTGHEURSPERLUG 7596  
| | | | |  
DB 7662 ACTCTCCAACAGCGAGTGACTGAGAACAGGGGCGGGGGGGGGGGCGGTGHEURSPERLUG 7711  
| | | | |  
QY 7597 TCAGGTCCAGCGAGACACAGGAGACG3CCTGCAAGAGCGGGGAGCTCCGACTCCCGC 7656  
| | | | |  
DB 7712 TCAGGTCCAGCGAGACACAGGAGACG3CCTGCAAGAGCGGGGAGCTCCGACTCCCGC 7771  
| | | | |  
QY 7657 AACCAAGAGAGAGCCCTGAGTCCGCGCTGCGCTTCATCATCTGTCTGCAAGAGCCG 7716  
| | | | |  
DB 7772 AACCAAGAGAGAGCCCTGAGTCCGCGCTGCGCGCTTCATCATCTGTCTGCAAGAGCCG 7831  
| | | | |  
QY 7717 GCATCTTGCCCTGCTTAAAGCCTTAACTTAAGCTCCCGCGGGGCTGGCCCTGTGCAAA 7776  
| | | | |  
DB 7832 GCATCTTGCCCTGCTTAAAGCCTTAACTTAAGCTCCCGCGGGGCTGGCCCTGTGCAAA 7891  
| | | | |  
QY 7777 CTTTACTCAGGGAGATGTTTACTGTGCTCGGAAAGGAGGAGGAGCGGAGCGG 7836  
| | | | |  
DB 7892 CTTTACTCAGGGAGATGTTTACTGTGCTCGGAAAGGAGGAGGAGCGGAGCGG 7951  
| | | | |  
QY 7837 GCAAGGCAAGCGGTGTGGGAGCAACAACAGCGGCGCAAGGCGGCAAGGAGCCCAAGACAG 7896  
| | | | |

DB 7952 GCACGCGAGCGGTGTGGGAGCCACACAGGCGGCGAGGCGGAGGAGCCCAAGACAG 8011  
| | | | |  
QY 7897 GATGACACACGACCTTCCACACGACCTGCGCTCCCGCAATGATTTTGAACCAAAAGCTTAA 7956  
| | | | |  
DB 8012 GATGACACACGACCTTCCACACGACCTGCGCTCCCGCAATGATTTTGAACCAAAAGCTTAA 8071  
| | | | |  
QY 7957 CTGAGCTGCGAGCCCGCGCGCTCCCTCCGCTTCCATCCCGCTTAAAGCGCTGTGAGCAG 8016  
| | | | |  
DB 8072 CTGAGCTGCGAGCCCGCGCGCTCCCTCCGCTTCCATCCCGCTTAAAGCGCTGTGAGCAG 8131  
| | | | |  
QY 8017 ATGAGCGAGGCGCTGTGCTCAAGCTCCCAAGTGGCTGTGTTCCGGTCCCCCAACAGCTGCCCA 8076  
| | | | |  
DB 8132 ATGAGCGAGGCGCTGTGCTCAAGCTCCCAAGTGGCTGTGTTCCGGTCCCCCAACAGCTGCCCA 8191  
| | | | |  
QY 8077 GCCAAGCAAGATTGCTGGAACCAAGTCAAGGCGAGGTGGGCGGACAAAGGGCCAGGTGG 8136  
| | | | |  
DB 8192 GCCAAGCAAGATTGCTGGAACCAAGTCAAGGCGAGGTGGGCGGACAAAGGGCCAGGTGG 8251  
| | | | |  
QY 8137 GCTTGGGGGGAACGAGTCTCCGAGACTGGAATGTTTTTTCACACATCGTTGCGCAG 8196  
| | | | |  
DB 8252 GCTTGGGGGGAACGAGTCTCCGAGACTGGAATGTTTTTTCACACATCGTTGCGCAG 8311  
| | | | |  
QY 8197 CGGTGGGAAGGAAGCGAGTGAATGATGTGTGTTTACAGGGTATATTTTGAATAC 8256  
| | | | |  
DB 8312 CGGTGGGAAGGAAGCGAGTGAATGATGTGTGTTTACAGGGTATATTTTGAATAC 8371  
| | | | |  
QY 8257 CTTCAATGAATTAATTGAGATGTTTTTACGCAAGAGACCTTACCAGATTAATGCTGTC 8316  
| | | | |  
DB 8372 CTTCAATGAATTAATTGAGATGTTTTTACGCAAGAGACCTTACCAGATTAATGCTGTC 8431  
| | | | |  
QY 8317 TGTGCTTTTGAATCTGTGTTTACCGTTCAAGAGCGGTGACAGGCCGACAGTGGTGAACC 8376  
| | | | |  
DB 8432 TGTGCTTTTGAATCTGTGTTTACCGTTCAAGAGCGGTGACAGGCCGACAGTGGTGAACC 8491  
| | | | |  
QY 8377 CATCATCTGCGAGACCAAGGGGGCGGGAGCTGCTGTGTAACGCGCCGCTGTGCTCCCTC 8436  
| | | | |  
DB 8492 CATCATCTGCGAGACCAAGGGGGCGGGAGCTGCTGTGTAACGCGCCGCTGTGCTCCCTC 8511  
| | | | |  
QY 8437 CTTCCCTTCCTTGGGCGAGATGAATTCAGATGCGTATCTGTGAGCGGCAATTTGCGCAGG 8496  
| | | | |  
DB 8552 CTTCCCTTCCTTGGGCGAGATGAATTCAGATGCGTATCTGTGAGCGGCAATTTGCGCAGG 8611  
| | | | |  
QY 8497 TGTGTAATCTGTGATTTACACAGCTGCTTAAATTAATAAGCAATTAATCTCAAAA 8556  
| | | | |  
DB 8612 TGTGTAATCTGTGATTTACACAGCTGCTTAAATTAATAAGCAATTAATCTCAAAA 8671  
| | | | |  
QY 8557 AAAAAAAAAAAAAA 8571  
| | | | |  
DB 8672 AAAAAAAAAAAAAA 8686  
| | | | |  
RESULT 3  
AR447713  
LOCUS AR447713 9053 bp DNA linear PAT 20-FEB-2004  
DEFINITION Sequence 306 from patent US 6673549.  
ACCESSION AR447713  
VERSION AR447713.1 GI:42676037  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 9053)  
AUTHORS Furness,L.M. and Buchbinder,J.L.  
TITLE Genes expressed in C3A liver cell cultures treated with steroids  
JOURNAL Patent: US 6673549-A 306 06-JUN-2004;  
Inocyte Corporation; Palo Alto, CA  
FEATURES  
source  
1..9053  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN



Query Match 96.1%; Score 8226.8; DB 6; Length 9053;  
Best Local Similarity 98.3%; Pred. No. 0;  
Matches 8451; Conservative 2; Mismatches 42; Indels 103; Gaps 10;

QY 1 CATGTGGGCTCCACACAGCTTGTGGACAGAGCTGAGGGCCACTGAGCCCCGCTAACCC 60  
DB 519 CATGTGGGATCCACACAGCTGTGGACAGAGTGGAGGGCCACTGAGCCCCGCTAACCC 578  
QY 61 GCGCCACAGCTTTTCTTACCCCATGAGATGCGCCCGGACGACACAGCACTCGGGCTCTCT 120  
DB 579 GCGCCACAGCTTTTCTTACCCCATGAGATGCGCCCGGACGACACAGCACTCGGGCTCTCT 638  
QY 121 GAGATCCACAGACCACTCCCGGACTATGCTCCACTGCTCGCGGGCTCCATCATCCA 180  
DB 639 GAGATCCACAGACCACTCCCGGACTATGCTCCACTGCTCGCGGGCTCCATCATCCA 698  
QY 181 GCGCCACGCGCGAGGCGCTCTCTGCTGTGAGTTCCAGCCCGGAAATGACGGTCCCA 240  
DB 699 GCGCCACGCGCGAGGCGCTCTCTGCTGTGAGTTCCAGCCCGGAAATGACGGTCCCA 758  
QY 241 GAGCTCCACCTGCGGCGACAGTCCCATCTATACCTGCGGAGCTGGGAAATCAGAGAT 300  
DB 759 GAGCTCCACCTGCGGCGACAGTCCCATCTATACCTGCGGAGCTGGGAAATCAGAGAT 818  
QY 301 GAGATTCAATTGAAAGCAAGCGCCTCGGCTAGAGCTGCTGACCCCTGCTGCGACC 360  
DB 819 GAGATTCAATTGAAAGCAAGCGCCTCGGCTAGAGCTGCTGACCCCTGCTGCGACC 878  
QY 361 GTTACCCCTGTGGCCACGCGGCGAGCTGCGGATCTGAGACCTGACCAAGACCGTAG 420  
DB 879 GTTACCCCTGTGGCCACGCGGCGAGCTGCGGATCTGAGACCTGACCAAGACCGTAG 938  
QY 421 CCTGACGGGCAAGCTGGAACCGGTGTCTCCCGGACCGCCGCACTGACCTTGAGCT 480  
DB 939 CCTGACGGGCAAGCTGGAACCGGTGTCTCCCGGACCGCCGCACTGACCTTGAGCT 998  
QY 481 GAGAGTGTGGCGGCAAGCTGTCTCAAGGAGAGCTGATCAGAACTGGAACCGCGTGA 540  
DB 999 GAGAGTGTGGCGGCAAGCTGTCTCAAGGAGAGCTGATCAGAACTGGAACCGCGTGA 1058  
QY 541 CCGAGAGATCACCATGTAGAGAGCAGACGATCTTAAAGCTGAAGAAAGACGACAGCT 600  
DB 1059 CCGAGAGATCACCATGTAGAGAGCAGACGATCTTAAAGCTGAAGAAAGACGACAGCT 1118  
QY 601 GAGAGAGAGAGCTGCCAAGCCGCCGAGCTTGAAAGCCCGTGTACCGCCGCCATCCA 660  
DB 1119 GAGAGAGAGAGCTGCCAAGCCGCCGAGCTTGAAAGCCCGTGTACCGCCGCCATCCA 1178  
QY 661 GTTGAAGCACCGGAGCTGTGTGAGATCATCTACGAGAAACCGGAAGAGGCTGAAGC 720  
DB 1179 GTTGAAGCACCGGAGCTGTGTGAGATCATCTACGAGAAACCGGAAGAGGCTGAAGC 1238  
QY 721 TGCACATCGGATTTCTGAAAGGCTGTGGGGCCCAAGTGAAGCTGCTGTACACAGGCC 780  
DB 1239 TGCACATCGGATTTCTGAAAGGCTGTGGGGCCCAAGTGAAGCTGCTGTACACAGGCC 1298  
QY 781 CTTCGACACCCCGGAGTATCATGAGAAATCAAAATTAACCAAGCGCATCGAAAGAGCT 840  
DB 1299 CTTCGACACCCCGGAGTATCATGAGAAATCAAAATTAACCAAGCGCATCGAAAGAGCT 1358  
QY 841 AATCTTGTACTTCAAGAGAGAAATCAAGCTCGGAAACAAATGAGAGAGATTCTGGCA 900  
DB 1359 AATCTTGTACTTCAAGAGAGAAATCAAGCTCGGAAACAAATGAGAGAGATTCTGGCA 1418  
QY 901 GCGCTATGACCAAGCTATGAGAGGCTTGGAAAAAAGGTGAGCGCATCGAAAAACAACC 960  
DB 1419 GCGCTATGACCAAGCTATGAGAGGCTTGGAAAAAAGGTGAGCGCATCGAAAAACAACC 1478  
QY 961 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGGAGTACTACGAAAGAGATTCCCTTGAGAT 1020  
DB 1479 GCGCGCGCGGCGCAAGAGAGCAAGGTGCGGAGTACTACGAAAGAGATTCCCTTGAGAT 1538  
QY 1021 CCGCAAGCAGCGGAGCTGACGAGAGCGCATGACAGAGAGGTTGGCAAGCGGCGAGTGG 1080

DB 1539 CCGCAAGCAGCGGAGCTGACGAGAGCGCATGACAGACAGGAGTGGCCAGCGGGCAGTGG 1598  
QY 1081 GCTGTCCATGTGGCCCGCCCGGACCGACACAGAGGTGTACGAGATCATGATGGCTCTTC 1140  
DB 1599 GCTGTCCATGTGGCCCGCCCGGACCGACACAGAGGTGTACGAGATCATGATGGCTCTTC 1658  
QY 1141 AGAGCAGAGAAACCTTGAGAGAGATGCGCGAGCTGGCGCGTATCCGCCCATGTGTA 1200  
DB 1659 AGAGCAGAGAAACCTTGAGAGAGATGCGCGAGCTGGCGCGTATCCGCCCATGTGTA 1718  
QY 1201 CGACGCTGACACAGCAGCGCATCAAGTTTCATCAATGAACGGGCTTATGCGCACCCCAT 1260  
DB 1719 CGACGCTGACACAGCAGCGCATCAAGTTTCATCAATGAACGGGCTTATGCGCACCCCAT 1778  
QY 1261 GAAAGTGTACAAAGACCCGCGAGCTATGAACATGTGAGTGTGACAGAGAAAGAGACTT 1320  
DB 1779 GAAAGTGTACAAAGACCCGCGAGCTATGAACATGTGAGTGTGACAGAGAAAGAGACTT 1838  
QY 1321 CCGGAGAAAGTTGATGACAGATCCCAAGAACTTGGCTGATGCAATCTTCTGAGAG 1380  
DB 1839 CCGGAGAAAGTTGATGACAGATCCCAAGAACTTGGCTGATGCAATCTTCTGAGAG 1898  
QY 1381 GAAAGCAGTGTGAGTGTGCTCTCTATTACTTACTGACTAAGAAATGAGAACTATAA 1440  
DB 1899 GAAAGCAGTGTGAGTGTGCTCTCTATTACTTACTGACTAAGAAATGAGAACTATAA 1958  
QY 1441 GAGCTGTGAGACGCGAGCTATGTGCGCGCGCAAGAGCTGAGACACACACACACAC 1497  
DB 1959 GAGCTGTGAGACGCGAGCTATGTGCGCGCGCAAGAGCTGAGACACACACACACAC 2018  
QY 1498 GCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1557  
DB 2019 GCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2078  
QY 1558 GAAAGATGAGAAAGAGAAAGAAAGAGCGGAGAGAGAGAGAGAGAGCGGAGGTGA 1617  
DB 2079 GAAAGATGAGAAAGAGAAAGAAAGAGAGCGGAGAGAGAGAGAGAGAGCGGAGGTGA 2138  
QY 1618 GAAACGACAAAGAAAGCTCTCTCAAGAGAGACAGACGACCTCAGGGAGAGACACGA 1677  
DB 2139 GAAACGACAAAGAAAGCTCTCTCAAGAGAGAGACAGACGACCTCAGGGAGAGACACGA 2198  
QY 1678 CAGAGAGAGGCTGTGGCTTCCAAAGCCGCAAAATGTGCAACAGCAAGGAGAGAGCGCA 1737  
DB 2199 CAGAGAGAGGCTGTGGCTTCCAAAGCCGCAAAATGTGCAACAGCAAGGAGAGAGCGCA 2258  
QY 1738 AGGCGCATCACCGCTCAATGGCTAATGAGGCCAACAGCGAGAGAGCCATCACCCCCA 1797  
DB 2259 AGGCGCATCACCGCTCAATGGCTAATGAGGCCAACAGCGAGAGAGCCATCACCCCCA 2318  
QY 1798 GCAAGCGCCGAGCTGCTTCCATGAGAGCTGAATGAGAGTTCTCGCTGACAGAGAAAG 1857  
DB 2319 GCAAGCGCCGAGCTGCTTCCATGAGAGCTGAATGAGAGTTCTCGCTGACAGAGAAAG 2378  
QY 1858 AATGAAACAGCCAAAGAAAGTCTCTGGAACAGCGCCGCAATGTGCGGCATTCGCCG 1917  
DB 2379 AATGAAACAGCCAAAGAAAGTCTCTGGAACAGCGCCGCAATGTGCGGCATTCGCCG 2438  
QY 1918 GATGTGGGCTCCAAAGCTGTGTGAGTGAAGAACTTCACTTCAACTCAAGAAAGAG 1977  
DB 2439 GATGTGGGCTCCAAAGCTGTGTGAGTGAAGAACTTCACTTCAACTCAAGAAAGAG 2498  
QY 1978 GCAAGAACTTGATGAGATTTTGACAGACACAACTGAAGATGAGAGAGAGAAACGC 2037  
DB 2499 GCAAGAACTTGATGAGATTTTGACAGACACAACTGAAGATGAGAGAGAGAAACGC 2558  
QY 2038 GCGAGAGAAAGAAAGAAAGCCCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 2096  
DB 2559 GCGAGAGAAAGAAAGAAAGCCCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2618  
QY 2097 TGGTGAAGATGAGAGATGAGAGCTCGGGCGTGAAGCGAAATGAGAGAGAGATGTGG 2156



Db 2619 TGGTGAAGATGAGAGATGAGAGCTCGGGCGTTGACCGGAAATGAGAGAGATGCTGG 2678  
QY 2157 AGAAGGCTGAAGCCTTACATGCTCTGGGAATGAGGTGCCCAGAGGGGAATGCAGTGGCC 2216  
Db 2679 AGAAGGCTGA----- 2688  
QY 2217 CAGCCACTGTCAACAAACAGCTCAAGACCCGAGAGCATCCCTCTCTCTCAACTGAGGCCG 2276  
Db 2689 -AGCCACTGTCAACAAACAGCTCAAGACCCGAGAGCATCCCTCTCTCTCAACTGAGGCCG 2747  
QY 2277 CCAAGGACACAGGGGAGAAATGGGCCCCAGCCCCCAGCCACCCCTGGGCGCCGAC-GGGGCA 2335  
Db 2748 CCAAGGACACAGGGGAGAAATGGGCCCCAGCCCCCAGCCACCCCTGGGCGCCGAGGGGCCA 2807  
QY 2336 CCCCCAGGCCACCAACCCCAACAACGAGAGACATCCCGGGCCCCCAATTGAGCCACCCCG 2395  
Db 2808 CCCCCAGGGGCCAACCCACCAACCGAGAGACATCCCGGGCCCCCACTGAGGCCAACCCCG 2867  
QY 2396 GCTCTGAAAGCCACCGGAGCCCTTACGCCCCCAACGACACCCCATGCTCTGTGACCT 2455  
Db 2868 GCTCTGAAAGCCACCGGAGCCCTTACGCCCCCAACGACACCCCATGCTCTGTGACCT 2927  
QY 2456 CTTCTGTGTGTCCCAAGGAGGAGAGAGAGAGAGACCGACGAGCGCCCCAGTGGAG 2515  
Db 2928 CTTCTGTGTGTCCCAAGGAGGAGAGAGAGAGAGACCGACGAGCGCCCCAGTGGAG 2987  
QY 2516 GAGGGGAGAGACAGAAAGCCCCCGCGCTGAGAGCTGGCAGTGGACACAGGGAAGGCC 2575  
Db 2988 GAGGGGAGAGACAGAAAGCCCCCGCGCTGAGAGCTGGCAGTGGACACAGGGAAGGCC 3047  
QY 2576 GAGGAGCCCTGTCAAGAGCAGTGCACAGAGAGAAACCGAGAGGGGCGGCCCAAGGGCAAG 2635  
Db 3048 GAGGAGCCCTGTCAAGAGCAGTGCACAGAGAGAAACCGAGAGGGGCGGCCCAAGGGCAAG 3107  
QY 2636 GACGCGAGAGCCGCTGAGAGCCACGCGGAGGGGGCGCTAAGGGCAAGAGAAAGAGGGCC 2695  
Db 3108 GACGCGAGAGCCGCTGAGAGCCACGCGGAGGGGGCGCTAAGGGCAAGAGAAAGAGGGCC 3167  
QY 2696 GGGAGCGGAGGGGCCACCACTGCAAGAGCTCGGGGCCCCCAGGACAGCGACCTCACT 2755  
Db 3168 GGGAGCGGAGGGGCCACCACTGCAAGAGCTCGGGGCCCCCAGGACAGCGACCTCACT 3227  
QY 2756 GCTTACTGTGATGACAGAGGTGATGAGGCCGAGGGCGGGCGACAAAGACCGGCTGCTG 2815  
Db 3228 GCTTACTGTGATGACAGAGGTGATGAGGCCGAGGGCGGGCGACAAAGACCGGCTGCTG 3287  
QY 2816 TCCCCAAGGCCAGGCTCTCAACCCGAGCTGGAGACCCCGGGGCAATGCTCAACCCAG 2875  
Db 3288 TCCCCAAGGCCAGGCTCTCAACCCGAGCTGGAGACCCCGGGGCAATGCTCAACCCAG 3347  
QY 2876 AAGCCACTGGAAGCTGAGAGAGCTGAGAGCAGAGCGGCTGCATCCGCCCAATCCAGTGC 2935  
Db 3348 AAGCCACTGGAAGCTGAGAGAGCTGAGAGCAGAGCGGCTGCATCCGCCCAATCC 3404  
QY 2936 ACAAAGTTCATGAGCCCCCGGGAGAGCAGACTCCCAACAAGCAGCTCCCGCCAG-C 2994  
Db 3405 ACAAAGTTCATGAGCCCCCGGGAGAGCAGACTCCCAACAAGCAGCTCCCGCCAGAC 3464  
QY 2995 CCCACCGGCAACGCAAAACCTGACGCGGAGAGCGAGCGGCCCTCAGCAGACCTGGACAGAG 3054  
Db 3465 CCCACCGGCAACGCAAAACCTGACGCGGAGAGCGAGCGGCCCTCAGCAGACCTGGACAGAG 3524  
QY 3055 CCCCCGGGCAAGAGAGAGGCCCGGACACCCCGCGCGCAAGAGAG----- 3100  
Db 3525 CCCCCGGGCAAGAGAGAGGCCCGGACACCCCGCGCGCAAGAGAGAGAGAGAGAGAGAGAGAG 3584  
QY 3101 -----GCTTTCGACAGCGAGGCCCAAGAGCTGCTGGGAGACCCCTTGTGAGAC 3150  
Db 3585 GTTCTTTCGACAGCTTCGACAGCGGCCCAAGAGCTGCTGGGAGACCCCTTGTGAGAC 3644  
QY 3151 TTTCCGCGCTGCGCTTCCCGGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3210  
Db 3645 TTTCCGCGCTGCGCTTCCCGGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3704

QY 3211 GGAACCTTCAGCTTCTCTCTACGCTTCACCTGGTCAACCACTGGCCCTTGGGCTTCATGA 3270  
Db 3705 GGAACCTTCAGCTTCTCTCTACGCTTCACCTGGTCAACCACTGGCCCTTGGGCTTCATGA 3764  
QY 3271 CATGCCCCGGGCGGCTCTGGCGGGCCCAACCCCAATCTCCAAACCGGCGCCCTCATCTC 3330  
Db 3765 CATGCCCCGGGCGGCTCTGGCGGGCCCAACCCCAATCTCCAAACCGGCGCCCTCATCTC 3824  
QY 3331 CTCTGCAAGACCCCAAGGCTCTGAGAGGCAAAATAGGTGCATCTCCAAAGAAATGTC 3390  
Db 3825 CTCTGCAAGACCCCAAGGCTCTGAGAGGCAAAATAGGTGCATCTCCAAAGAAATGTC 3884  
QY 3391 GGTTCAGCTTCAGCTTCCGTAATCAAGAGATGCCAAGGCCCGGTGGGCTGTCAAT 3450  
Db 3885 GGTTCAGCTTCAGCTTCCGTAATCAAGAGATGCCAAGGCCCGGTGGGCTGTCAAT 3944  
QY 3451 GGGGCTGCCCCCTGCCCCAATGAGACCCCAAAAGTGGACACCTTCAACCGAGTGAAGAGGA 3510  
Db 3945 GGGGCTGCCCCCTGCCCCAATGAGACCCCAAAAGTGGACACCTTCAACCGAGTGAAGAGGA 4004  
QY 3511 GCAGCTGTCCCAACGGGACAGGCTGGGCAACCGAGAGACCTGGGGGTGCCACAGCCA 3570  
Db 4005 GCAGCTGTCCCAACGGGACAGGCTGGGCAACCGAGAGACCTGGGGGTGCCACAGCCA 4064  
QY 3571 GGAAGGCTTCGCTGTGAGAGGGAAGCTTGGGCTCAAGTTCGGGCGGAAGCATCA 3630  
Db 4065 GGAAGGCTTCGCTGTGAGAGGGAAGCTTGGGCTCAAGTTCGGGCGGAAGCATCA 4124  
QY 3631 AGGCATTCGAGCAACAGGGTGCCTCGGACAGGGCCATACATACCGGGCTCATAC 3690  
Db 4125 AGGCATTCGAGCAACAGGGTGCCTCGGACAGGGCCATACATACCGGGCTCATAC 4184  
QY 3691 CCACGGCAACGCAAGCTGACGCTCTGTACAGGGCAACATCACAGAGATATCGGAGGA 3750  
Db 4185 CCACGGCAACGCAAGCTGACGCTCTGTGTACAGGGCAACATCACAGAGATATCGGAGGA 4244  
QY 3751 CAGCCGAGTGGCTTGAACCGCGGCGGAGAGACAGCTTGCCCAAGGGCACGTCACTTA 3810  
Db 4245 CAGCCGAGTGGCTTGAACCGCGGCGGAGAGACAGCTTGCCCAAGGGCACAGTCACTTA 4304  
QY 3811 CGAAGCAAGAAAGGGCCACGCTTGTCTTATGAGGGTGGCAATGTGTGACCCCAATGCTC 3870  
Db 4305 CGAAGCAAGAAAGGGCCACGCTTGTCTTATGAGGGTGGCAATGTGTGACCCCAATGCTC 4364  
QY 3871 CAAGAGACGCGCAAGAGCAAGCTCAGGACCCCCCATGAGACGGGCGGCCCAAGCGCAC 3930  
Db 4365 CAAGAGACGCGCAAGAGCAAGCTCAGGACCCCCCATGAGACGGGCGGCCCAAGCGCAC 4424  
QY 3931 CTATGACATGATGAGAGGCGCGGTGGAGAGCAATCTTCAAGCCAGCATTCAGAGTCT 3990  
Db 4425 CTATGACATGATGAGAGGCGCGGTGGAGAGCAATCTTCAAGCCAGCATTCAGAGTCT 4484  
QY 3991 CATGGGCGGTGCATTCGCGCGGAGAGCAACAGCCCCCAACCTTCAAGAGACACACA 4050  
Db 4485 CATGGGCGGTGCATTCGCGCGGAGAGCAACAGCCCCCAACCTTCAAGAGACACACA 4544  
QY 4051 CATCGCGGGTGCATTCACACAAGAGGATCCTTGATCTTACGTTGAGAGCAAGAGAGACTTA 4110  
Db 4545 CATCGCGGGTGCATTCACACAAGAGGATCCTTGATCTTACGTTGAGAGCAAGAGAGACTTA 4604  
QY 4111 CTTGCTGTGGGAGGCAAGCTCTTAAGCGGAGGGGACGCTTCGCCCCACGCGCTTC 4170  
Db 4605 CTTGCTGTGGGAGGCAAGCTCTTAAGCGGAGGGGACGCTTCGCCCCACGCGCTTC 4664  
QY 4171 ACGGACCTGACCGAGGCTTACAGAGCGAGGCTCTGGGCGCCCTTGAAGTGAAGCGGCG 4230  
Db 4665 ACGGACCTGACCGAGGCTTACAGAGCGAGGCTCTGGGCGCCCTTGAAGTGAAGCGGCG 4724  
QY 4231 CCAATGAGGCGCTGAGGCAAGGTAAGAGAGCGGGCGGCTCATCTAGATCCGCG 4290  
Db 4725 CCAATGAGGCGCTGAGGCAAGGTAAGAGAGCGGGCGGCTCATCTAGATCCGCG 4784

QY	4291	CGAGAGCTGGGGCAACAGCCCGAGCTGCTTGGCCCGGGGGCCGCTCAAGAGAGGGCTC	4350
Db	4785	CGAGAGCTGGGGCAAGCCCGAGCTGCTTGGCCCGGGGGCCGCTCAAGAGAGGGCTC	4844
QY	4351	CATCAGCAGAGGGACCCCGCTCAAGTACGACACGGGCGCTGCACACATGGCTCCAAAA	4410
Db	4845	CATCAGCAGAGGGACCCCGCTCAAGTACGACACGGGCGCTGCACACATGGCTCCAAAA	4904
QY	4411	GCACGACGTACGCTCCCTCATCGGCAAGCCCGGCGGAGCGTTCCACCCCGTGACCCGCT	4470
Db	4905	GCACGACGTACGCTCCCTCATCGGCAAGCCCGGCGGAGCGTTTCCACCCCGTGACCCGCT	4964
QY	4471	GGATGTGATGGCCCGACCGCCCGGGCACTGAAACGTGCTGCTACAGAGAGAGCTGAAGAG	4530
Db	4965	GGATGTGATGGCCCGACCGCCCGGGCACTGAAACGTGCTGCTACAGAGAGAGCTGAAGAG	5024
QY	4531	CCGGCAGAGGACCGGCAAGCAGCTCGGGGGGGCTCCATTGGCGGCGGCGCCCGGCTATTGT	4590
Db	5025	CCGGCAGAGGACCGGCAAGCAGCTCGGGGGGGCTCCATTGGCGGCGGCGCCCGGCTATTGT	5084
QY	4591	GCCTGAGCTGGGTAAAGCCGCGGCAAGAGCCGCCCTTGAACCTATGAGGACCACGGGGACCCCT	4650
Db	5085	GCCTGAGCTGGGTAAAGCCGCGGCAAGAGCCGCCCTTGAACCTATGAGGACCACGGGGACCCCT	5144
QY	4651	TGCGGGCCACCTCCCAAGAGGTGGCCCGTGTACAATGCGGGAGCCCAAGCCGCCCTTGCA	4710
Db	5145	TGCGGGCCACCTCCCAAGAGGTGGCCCGTGTACAATGCGGGAGCCCAAGCCGCCCTTGCA	5204
QY	4711	GGAGGGGACGCTTTCGTCCAGCAGAGGACATCCCAAGACCGAAAGGTGACGTGAGCGCTCG	4770
Db	5205	GGAGGGGACGCTTTCGTCCAGCAGAGGACATCCCAAGACCGAAAGGTGACGTGAGCGCTCG	5264
QY	4771	TGAGATGCGCAGTGTCCCGGCAACAGCCGTCGCCGAGACCAACCCACACCCCATCTGACC	4830
Db	5265	TGAGATGCGCAGTGTCCCGGCAACAGCCGTCGCCGAGAGCAACCCACACCCCATCTGACC	5324
QY	4831	CTATGAGCACCCTGTGGGGCGGTGAGTGGCGTGAACCTGTATGCGACGACATCCCTCT	4890
Db	5325	CTATGAGCACCCTGTGTGGGGCGGTGAGTGGCGTGAACCTGTATGCGACGACATCCCTCT	5384
QY	4891	GGCCTTGGACCCCAACCTCCATACCCCGCGGATCCCTTGGAGCGACGCGGTGCTACTA	4950
Db	5385	GGCCTTGGACCCCAACCTCCATACCCCGCGGATCCCTTGGAGCGACGCGGTGCTACTA	5444
QY	4951	CCTGCCCCGACACCTGGGCCCCCAACCCCACTAACCGGACCTGTACCCACCCCTACTCAT	5010
Db	5445	CCTGCCCCGACACCTGGGCCCCCAACCCCACTAACCGGACCTGTACCCACCCCTACTCAT	5504
QY	5011	CCGCGGCTACCCCGACACGAGGCGGCGTGTGAGAACCGGACGACATCAATGAATGACTCAT	5070
Db	5505	CCGCGGCTACCCCGACACGAGGCGGCGTGTGAGAACCGGACGACATCAATGAATGACTCAT	5564
QY	5071	CACCTTCGACAGATGCAACCAACAGGAGCAACCGGCAATGGCCCAAGGAGGTGATATGCT	5130
Db	5565	CACCTTCGACAGATGCAACCAACAGGAGCAACCGGCAATGGAGGTGATATGCT	5624
QY	5131	GAGGGGACCTTCGCGCCCGGAGTGTCTGTGGCTGGACTCAATCAACGTGCGGGTCCCGAGG	5190
Db	5625	GAGGGGACCTTCGCGCCCGGAGTGTCTGTGGCTGGACTCAATCAACGTGCGGGTCCCGAGG	5684
QY	5191	CATCATGACCTGTGCCAAGTGCACACCTTGCTGTGTGCTGCGGCCCGGACACGAGGCAC	5250
Db	5685	CATCATGACCTGTGCCAAGTGCACACCTTGCTGTGTGCTGCGGCCCGGACACGAGGCAC	5744
QY	5251	CCGAGGCCACCGGCAATGACCGGCTTTGGCTACCTGCCACCGGCGGCCGACGCTTCAAGCAG	5310
Db	5745	CCGAGGCCACCGGCAATGACCGGCTTTGGCTACCTGCCACCGGCGGCCGACGCTTCAAGCAG	5804
QY	5311	CCGCGCACAGACGTCCCCCACTCTCCCGGAGAGTTCACACACTTGAACAAACGAC	5370
Db	5805	CCGCGCACAGACGTCCCCCACTCTCCCGGAGAGTTCACACACTTGAACAAACGAC	5864
QY	5371	CACGCTCTCGTCCGAGCGGAGCGAGACCGGGAATCGAGAGCGGAGCCGGATTCGAGAGCG	5430

Db	5865	CACGTCCTCGTCCGAGCGGGAGCGAGACCCGGGATCGAGAGCGGGACCGGGATCCGGAGCG	5924
QY	5431	GGAAAGTCATCTCTCAAGTCACACGACCGGTGAGACACGACCCATCTGAGAACCTTGG	5490
Db	5925	GGAAAGTCATCTCTCAAGTCACACGACCGGTGAGACACGACCCATCTGAGAACCTTGG	5984
QY	5491	TACGAGCAGAGCAGCGCGCAGACCGGCAAGACGCGCGGGGTGTGGGGGCAGACACACCG	5550
Db	5985	TACGAGCAGAGCAGCGCGCAGACCGGCAAGACGCGCGGGGTGTGGGGGCAGACACACCG	6044
QY	5551	CCCGGCGTCCCACTCCCATGCGCCACACGACACTGCGCCATCTCCCTCGGAGCCGAGATGC	5610
Db	6045	CCCGGCGTCCCACTCCCATGCGCCACACGACACTGCGCCATCTCCCTCGGAGCCGAGATGC	6104
QY	5611	CCTCCAGCAGAGACCCAGTGTGCTTCAACAACAGGATGAGGGATTCATCAACCGGTGT	5670
Db	6105	CCTCCAGCAGAGACCCAGTGTGCTTCAACAACAGGATGAGGGATTCATCAACCGGTGT	6164
QY	5671	GGAGCCGACGACGAGCCCAAGTCTTG-----AGTCCACTCTCAACTCTCTCAACCCGT	5721
Db	6165	GGAGCCGACGACGAGCCCAAGTCTTG-----AGTCCACTCTCAACTCTCTCAACCCGT	6224
QY	5722	TGCGCCAGCTGCGCATTTCCCACTGTGACACCACTGCGCACTGGGGCGGCACTCCGATGG	5781
Db	6225	TGCGCCCGCTGCGCATTTCCCACTGTGACACCACTGCGCACTGGGGCGGCACTCCGATGG	6284
QY	5782	GGTCTACCCCTTACCCCTCATGAGAGCCGTCCTTGCTGCGCCAGAGAGGCCCGCGGATCGCG	5841
Db	6285	GGTCTACCCCTTACCCCTCATGAGAGCCGTCCTTGCTGCGCCAGAGAGGCCCGCGGATCGCG	6344
QY	5842	GCCAGAGCGGCGCCCGAGCAGACACCGGCAATGCTTCTTGCCAGAGCCCAAGCCCGCTC	5901
Db	6345	GCCAGAGCGGCGCCCGAGCAGACACCGGCAATGCTTCTTGCCAGAGCCCAAGCCCGCTC	6404
QY	5902	CGGCGTGAAGCGCGGCTCTCTCCCAACAGAGGCTCGGAGCGCGCGGCGCCCTGAGGCGTCC	5961
Db	6405	CGGCGTGAAGCGCGGCTCTCTCCCAACAGAGGCTCGGAGCGCGCGGCGCCCTGAGGCGTCC	6464
QY	5962	TGTCTTGAGCACGACACATCGCGCGCACCCCTGCGAAGAACCTTCGCACTTCAACAGC	6021
Db	6465	TGTCTTGAGCACGACACATCGCGCGCACCCCTGCGAAGAACCTTCCGCACTTCAACAGC	6524
QY	6022	CAGCCCGGACCCCGCGCGCCACTTGTCTTGCGCTCGGACCCCGACCGGGAAAGACTCA	6081
Db	6525	CAGCCCGGACCCCGCGCGCCACTTGTCTTGCGCTCGGACCCCGACCGGGAAAGACTCA	6584
QY	6082	AAGTAAACCTTTTCCATCCAGGAACTGGAATCTCCGTTCTTGCGTTAACACCGGCAACG	6141
Db	6585	AAGTAAACCTTTTCCATCCAGGAACTGGAATCTCCGTTCTTGCGTTAACACCGGCAACG	6644
QY	6142	CTACAGCCCGGAAGGGGTGAGCGCGTCAAGCGCTGTAGACTCACCGATCTGACCCACGA	6201
Db	6645	CTACAGCCCGGAAGGGGTGAGCGCGTCAAGCGCTGTAGACTCACCGATCTGACCCACGA	6704
QY	6202	CAGGCGGCTCCCAAGACCTTGAAGACTTGAACAAGGCCACTTGAAGGGAGGACTGCG	6261
Db	6705	CAGGCGGCTCCCAAGACCTTGAAGACTTGAACAAGGCCACTTGAAGGGAGGACTGCG	6764
QY	6262	GCCCAAGCAGCAGGCGCCGTTGAAGCTTGGGCGGGGAAGCGGCGCACTTCCACACTTGG	6321
Db	6765	GCCCAAGCAGCAGGCGCCGTTGAAGCTTGGGCGGGGAAGCGGCGCACTTCCACACTTGG	6824
QY	6322	GCGCTGCTGAGAGGACGAGCCCTGTCAAGCGCGCTCTCAAGCGCGCCCAAGGGGTCAA	6381
Db	6825	GCGCTGCTGAGAGGACGAGCCCTGTCAAGCGCGCTCTCAAGCGCGCCCAAGGGGTCAA	6884
QY	6382	AGGTCAACGAGCGGCTGTCACTCTGCGCCAGACATCATGTGAGGTTCATCAACAGACTA	6441
Db	6885	AGGTCAACGAGCGGCTGTCACTCTGCGCCAGACATCATGTGAGGTTCATCAACAGACTA	6944
QY	6442	CACCCGACACACCCACAGACGCTCAAGGCACTTGCCTGCGCGGCGCCCTCTTACTCTTCC	6501

D	645	CA	CCCGG	CA	CCCA	CCCA	CAG	CA	AG	CT	CAG	GC	AC	CC	CT	GC	CC	CC	CC	CT	CT	CT	CT	CC	7004						
Q	6502	TG	GGG	CA	CA	CT	GC	CC	CT	CT	GA	CT	CC	GC	CC	CA	CC	CA	GT	GA	CT	CT	CC	GC	6561						
D	7005	TG	GGG	CA	CA	CT	GC	CC	CT	CT	GA	CT	CC	GC	CC	CA	CC	CA	GT	GA	CT	CT	CC	GC	7064						
Q	6562	CC	CGG	CA	CA	TG	GT	AC	CC	CG	CC	CT	GC	CT	CC	CA	CA	GC	GA	GG	GG	CA	GA	GT	CT	CC	6621				
D	7065	CC	CGG	CA	CA	TG	GT	AC	CC	CG	CC	CT	GC	CT	CC	CA	CA	GC	GA	GG	GG	CA	GA	GT	CT	CC	7124				
Q	6622	AG	AG	CA	AA	CA	GA	AG	CT	GC	CT	TG	GG	GT	GT	GA	GA	CG	GT	AT	GA	AA	CG	GT	CT	CC	6681				
D	7125	AG	AG	CA	AA	CA	GA	AG	CT	GC	CT	TG	GG	GT	GT	GA	GA	CG	GT	AT	GA	AA	CG	GT	CT	CC	7184				
Q	6682	AC	CGA	GG	CG	CA	TGA	CG	GA	CC	GA	GG	CA	CT	CC	CG	AG	TC	GT	GT	AT	CA	CC	CG	TC	GT	AC	6741			
D	7185	AC	CGA	GG	CG	CA	TGA	CG	GA	CC	GA	GG	CA	CT	CC	CG	AG	TC	GT	GT	AT	CA	CC	CG	TC	GT	AC	7244			
Q	6742	GG	AT	GG	GG	GA	CA	GA	CG	AG	CC	CA	GG	AT	GG	GT	CC	CA	AG	TC	CA	GG	CA	CA	CA	GG	CA	6801			
D	7245	GG	AT	GG	GG	GA	CA	GA	CG	AG	CC	CA	GG	AT	GG	GT	CC	CA	AG	TC	CA	GG	CA	CA	CA	GG	CA	7304			
Q	6802	GC	CG	CA	GG	CC	CT	CT	CT	CA	GA	CA	GT	CA	CC	GA	AG	CA	CT	CC	CG	CA	TG	GT	CA	AG	TC	CA	AG	AA	6861
D	7305	GC	CG	CA	GG	CC	CT	CT	CT	CA	GA	CA	GT	CA	CC	GA	AG	CA	CT	CC	CG	CA	TG	GT	CA	AG	TC	CA	AG	AA	7364
Q	6862	GCA	AG	AT	CA	CA	GA	AG	CT	GA	CA	CC	CA	CA	CC	GA	AT	GA	AG	CC	CT	GA	TA	CA	AT	CA	G	6921			
D	7365	GCA	AG	AT	CA	CA	GA	AG	CT	GA	CA	CC	CA	CA	CC	GA	AT	GA	AG	CC	CT	GA	TA	CA	AT	CA	G	7424			
Q	6922	CCA	GC	CT	GG	CA	CG	GA	AT	CT	CT	CA	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	6981				
D	7425	CCA	GC	CT	GG	CA	CG	GA	AT	CT	CT	CA	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT	7484				
Q	6982	TAGA	AG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	7041		
D	7485	TAGA	AG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	GG	CA	7544		
Q	7042	AAA	GG	CA	CT	CA	TG	GG	GT	AA	TA	TG	CA	CA	TG	GG	AA	AG	TC	CC	CG	CG	CT	CA	CG	CG	CA	AT	GC	7101	
D	7545	AAA	GG	CA	CT	CA	TG	GG	GT	AA	TA	TG	CA	CA	TG	GG	AA	AG	TC	CC	CG	CG	CT	CA	CG	CG	CA	AT	GC	7604	
Q	7102	TTTT	AA	CC	CT	CT	GA	AT	GC	CA	GT	GC	CA	GC	CT	GC	CT	GC	CT	GC	CT	GC	CT	GC	CT	GC	CT	7161			
D	7605	TTTT	AA	CC	CT	CT	GA	AT	GC	CA	GT	GC	CA	GC	CT	GC	CT	GC	CT	GC	CT	GC	CT	GC	CT	GC	CT	7161			
Q	7162	CG	GA	CG	GA	GT	CA	CA	CA	CT	CA	CT	CG	GC	CA	GT	GG	CG	GG	GA	AG	GC	CA	AG	TC	CT	GC	7221			
D	7665	CG	GA	CG	GA	GT	CA	CA	CA	CT	CA	CT	CG	GC	CA	GT	GG	CG	GG	GA	AG	GC	CA	AG	TC	CT	GC	7221			
Q	7222	CAGA	CC	CA	GG	CA	CG	CG	AA	AG	CA	GA	TC	CC	CG	GC	CC	CG	GC	CT	GC	CA	TG	GG	GG	GA	CG	GC	7281		
D	7725	CAGA	CC	CA	GG	CA	CG	CG	AA	AG	CA	GA	TC	CC	CG	GC	CC	CG	GC	CT	GC	CA	TG	GG	GG	GA	CG	GC	7281		
Q	7282	ACC	CT	CT	GT	CT	CT	CA	GT	GA	CA	CT	CG	GA	GG	GA	GT	GA	CT	GA	CC	GC	CG	GA	CG	GC	CT	CA	CG	7341	
D	7785	ACC	CT	CT	GT	CT	CT	CA	GT	GA	CA	CT	CG	GA	GG	GA	GT	GA	CT	GA	CC	GC	CG	GA	CG	GC	CT	CA	CG	7341	
Q	7342	CC	GG	CT	GG	GA	GG	GA	CA	AG	CC	CT	CG	TC	CG	GA	GT	TC	CA	CC	CA	AT	CC	CC	CT	CA	CC	CC	CT	7401	
D	7845	CC	GG	CT	GG	GA	GG	GA	CA	AG	CC	CT	CG	TC	CG	GA	GT	TC	CA	CC	CA	AT	CC	CC	CT	CA	CC	CC	CT	7401	
Q	7402	GAT	CAT	GC	GG	CT	GC	AG	CG	GG	GT	CA	TG	CA	TG	CC	CA	CC	CC	CA	CG	GG	CT	CC	CG	CG	CG	CG	7461		
D	7905	GAT	CAT	GC	GG	CT	GC	AG	CG	GG	GT	CA	TG	CA	TG	CC	CA	CC	CC	CA	CG	GG	CT	CC	CG	CG	CG	CG	7461		
Q	7462	CAG	CG	GG	CC	CT	GC	TG	GC	CC	CC	CA	CA	CG	CT	GG	GA	CG	AG	GC	CA	GA	CG	CA	GA	CG	CA	GA	7521		
D	7965	CAG	CG	GG	CC	CT	GC	TG	GC	CC	CC	CA	CA	CG	CT	GG	GA	CG	AG	GC	CA	GA	CG	CA	GA	CG	CA	GA	7521		
Q	7522	CT	CG	CA	GT	AC	GA	GA	CA	CT	CT	CC	GA	CA	GG	AG	AT	GA	CT	CA	GA	AA	CA	GG	GG	GG	GG	GG	7581		
D	8025	CT	CG	CA	GT	AC	GA	GA	CA	CT	CT	CC	GA	CA	GG	AG	AT	GA	CT	CA	GA	AA	CA	GG	GG	GG	GG	GG	7581		

OY		7582	CGGTBERSPBRUJGTGAGGTCCCAAGCGAACACAGGAACGGGCCCTTCAGAGAATCGGGGC	7641
Dp	:	8085	-----GCGGTGTACAGGTCCAGGAGCACAGGAAGGGCCTTCAGAGAATCGGGGC	8135
OY		7642	GCTGCAGACTCCCCAAACCAAAGGAAGAGGCCCTTGAGTCCGCTCGCATCTCATCTCT	7701
Dp		8136	GCTGCAGACTCCCCAAACCAAAGGAAGAGGCCCTTGAGTCCGCTCGCATCTCATCTCT	8195
OY		7702	GTCGCTCCAGAGCGCGGANTCTTTGCTGTATAAGCTTAAGATTCTCCGCCC	7761
Dp		8196	GTCGCTCCAGAGCGCGGANTCTTTGCTGTATAAGCTTAAGATTCTCCGCCC	8255
OY		7762	CTGGCCCTGTGTCAGACTTTACTAGGGGATGTTTACTGTGCTCGGGAAGGAGGGGAA	7821
Dp		8256	CTGGCCCTGTGTCAGACTTTACTAGGGGATGTTTACTGTGCTCGGGAAGGAGGGGAA	8315
OY		7822	GGGGCCGGGGAGGGGGGACGGCAGGCGTGTGGACACAACAAGGCGGCACGGGCGGC	7881
Dp		8316	GGGGCCGGGGAGGGGGGACGGCAGGCGTGTGGACACAACAAGGCGGCACGGGCGGC	8375
OY		7882	AGGGACCCAAGACGAGTAGACACAAGCATCTCAAGCAGCTGCTCCCGAATGATTTG	7941
Dp		8376	AGGGACCCAAGACGAGTAGACACAAGCATCTCAAGCAGCTGCTCCCGAATGATTTG	8435
OY		7942	GAACCAAAAGTCTAACTGAGCTGCGACGCCCGGCTTCCTCCGCTCCCATCCGCT	8001
Dp		8436	GAACCAAAAGTCTAACTGAGCTGCGACGCCCGGCTTCCTCCGCTCCCATCCGCT	8495
OY		8002	TAGGCGCTGTGACAGATGAGACGACGCCCTGTCCAGGCCCAAGTGGCTGTCCGCTTC	8061
Dp		8496	TAGGCGCTGTGACAGATGAGACGACGCCCTGTCCAGGCCCAAGTGGCTGTCCGCTTC	8555
OY		8062	CCAAGACTGCTCCCGACCAACGAGATTGCTGGAACCAATGCAGGCCAGGTGGCGGACA	8121
Dp		8556	CCAAGACTGCTCCCGACCAACGAGATTGCTGGAACCAATGCAGGCCAGGTGGCGGACA	8615
OY		8122	AAAAGGCGCAGGTGCGGCTTGGGGGAAACGATGCTCCAGAGACTGACTGTTTTTTTAC	8181
Dp		8616	AAAAGGCGCAGGTGCGGCTTGGGGGAAACGATGCTCCAGAGACTGACTGTTTTTTTAC	8675
OY		8182	ACATCGTTGCGCGACGCGTGGGAAAGGAAGGAGTAAATGATGTGTTTACAG	8241
Dp		8676	ACATCGTTGCGCGACGCGTGGGAAAGGAAGGAGTAAATGATGTGTTTACAG	8735
OY		8242	GTAATATTTTGATPACTCTTCATGATTAATTAATTCAGATTTTTCAGCAGAAGACTTACC	8301
Dp		8736	GTAATATTTTGATPACTCTTCATGATTAATTAATTCAGATTTTTCAGCAGAAGACTTACC	8795
OY		8302	CAGTATTAATGCTGCTGCTGTGCTTTTGATCTGTGCTTACCGTTCAAGAGCGTGTGCAGGCC	8361
Dp		8796	CAGTATTAATGCTGCTGCTGTGCTTTTGATCTGTGCTTACCGTTCAAGAGCGTGTGCAGGCC	8855
OY		8362	GACAGTGTGTAAGCCCACTCACTGCGCAGAGAACMAAGGGGGCGGGAGTCTCG-TCACGCC	8420
Dp		8856	GACAGTGTGTAAGCCCACTCACTGCGCAGAGAACMAAGGGGGCGGGAGTCTCGTCAACGCC	8915
OY		8421	CGCGTGTCCCTCCCTCCCTCCCTCTTCTGTGGCAGAAATGAAATTCGATTCGTGTGCG	8480
Dp		8916	CGCGTGTCCCTCCCTCCCTCCCTCTTCTGTGGCAGAAATGAAATTCGATTCGTGTGCG	8975
OY		8481	CGCATTTTGGCAGAGGTGATGTAATTCGTCAATTTACACAGTGGTTCTAATTTAAAAAGC	8540
Dp		8976	CGCATTTTGGCAGAGGTGATGTAATTCGTCAATTTACACAGTGGTTCTAATTTAAAAAGC	9035
OY		8541	GAATTATATCTCAAAAAA 8558       	
Dp		9036	GAATTATATCTCAAAAAA 9053       	

RESULT\_4  
 AY965853      8548 bp    mRNA    linear    PRI 11-Apr-2005

DEFINITION	Accession
Human sapiens SMRTe-tau mRNA, complete cds, alternatively spliced.	AY65853
Human sapiens SMRTe-tau mRNA, complete cds, alternatively spliced.	AY65853.1 GI:62240097
ORGANISM	Homo sapiens (human)
KEYWORDS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	Chen, J.D. and Evans, R.M. A transcriptional co-repressor that interacts with nuclear hormone receptors
AUTHORS	Nature 377 (6548), 454-457 (1995)
JOURNAL	7566127
PUBMED	2 (bases 1 to 8548)
REFERENCE	Park, E.J., Schreier, D.J., Yang, M., Li, H., Li, L. and Chen, J.D. SMRE, a silencing mediator for retinoid and thyroid hormone receptors-extended isoform that is more related to the nuclear receptor corepressor
AUTHORS	Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3519-3524 (1999)
JOURNAL	10097068
PUBMED	3 (bases 1 to 8548)
REFERENCE	Li, C.-W., Johnson, D., Chen, L.-Y., Ghosh, J. and Chen, J.D. Regulation and binding of the steroids and xenobiotics receptor SKS/PXR by nuclear receptor corepressor SMRE
AUTHORS	Unpublished
JOURNAL	4 (bases 1 to 8548)
PUBMED	Chen, J.D.
REFERENCE	Direct Submission
AUTHORS	Submitted (17-MAR-2005) Pharmacology, UMDN7-Robert Wood Johnson Medical School, 661 Hoes Lane, Piscataway, NJ 08901, USA
JOURNAL	Location/Qualifiers
FEATURES	1..8548
source	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	157..7542
	/note="silencing mediator of retinoid and thyroid hormone receptors extended isoform splice variant 1; transcriptional corepressor; SMRE; alternatively spliced"
	/codon_start=1
	/product="SMRE-tau"
	/protein_id="AAI77219.1"
	/db_xref="GI:62240098"
	/translation="MSGSTOPVQWTRATBRYPHSLSYVQIARTHTDVGLEVOHHSRYRHLSLSPGSIIOBORRRLPSIFROPENRSOHLAPRSHVYPLIGSSEMPITBSRRPLLELPPLLRPSILATGCPAGSBDLTKDSLTGKTEPVPSPSPHTDPELIVPRPLSKELLIONMRVDEITVMVQISLKKQOOLEEKAPPEPEKVPSPPIIESGHSILVQIYDENRKKAAHRIIEGLQVLELPYNOPTSTROYHENIKINQAMRKILLYFRNRNARKOMEQFCQYDOLMAWKEKTERLENNRPAKSKSKVEYEEKOPEIKKORELOERMORVQGGSGLSMSRASHVEYSEIIDGLSDENLEKOROLAVIPMLYDADQQRKIFTNMGLNADPMKYTVKDRVNMMSQEQETREKTMQNPXFGILASFLSKRTVAECVLITYITKKNENYKSLVRSYTRKGSQOQOQOQOQOQOQOQOQMPRRSQSEERDEKEKEKEKEKEKEKEKEKEKEKEKEKEDDLKEKTDTSGBDNDEKAVAKGRKTANSQGRKGRITRSMANENASEEALTPQOQALIASLNNSSSWTEENMTAKKRLGLEHNMWSAIAMVGSKTVSCKNFYVYKKQNDLILQKLMKEGRNARKKKAKKAPNANWSERAPPVVEDEMEASGVTAENNEVEEAAATVNNSSPTESLPSHTAIAKTCQNGCRPRATLGADGPPGCPPTPEPPIIPATPSTASATLTATPPAPPPSPPEPVAEKEKEEETAPAPVEEGEBOIPRAEELAVDTGAEIPVSECTEAEABEGAKGDAAAEKTEARALKAERKEGGSRAATTAKSGAPQDSSTASABEVABEABGDKRLSLRPSRLPTPTGDPANAPAPQRPDLKQLKORAAAIPIQVTKVHPREEDAAPTKRAPAPPPPPONLQESDAPAPQPGSSPRKSSPAPPADKEAEKVPFAAALAAQKLDEDPCCMTSGLPFPVPREVIVKASHPADPAPSAFVAPGHPPLGLHTARVILBRPPTSNRPPLISAKHPSVLEQIGASQCSQSVQHLHYPSHAAAPGVPYTMGLPLMDPKKAPFGSGVQQLSPRGQGPSPESGVPTAEASVLEKTAGASVPGSITKGASITRVSDSAITVRSITHTGTPADVLVYKTTIRIIGEDSPSLRDRERDSLPRGHVYEGKTVSHVLSIEGMSVTOCSKEDGSSSPPHETAPKTYDMEERVERALISASIEELMGRIAPERSPHLKEOHNRGSIITGOIPRSYVEAQDYLREAKLIKEGTPEPPPSRLATVAKYQALGPKIKRAHGALVATVEARSIHETPRELHTPPLPLAPRPLKEGILNCTGTPLYKVTGASTTSGSKHDVRLISGSGRTFPPHPLDWMADRALERACTBSLKSGTGHSSSGSIARCAPITVELGKPSPLTYEDHGAPFGHILPRGSPVPMREPPRLQEGSLSSSKASQARKLTSTPRELAKPSSTVPEHHPPPISEPHLLKVGSGV

Db	156	CATGTCGGGGCTCCACACAGAGCTTGTGGACAAGACGTGGAGGGCCACTAGAGCCGCCCTACCC	60	1	CATGTCGGGGCTCCACACAGAGCTTGTGGACAAGACGTGGAGGGCCACTAGAGCCGCCCTACCC	60
Qy	1	CATGTCGGGGCTCCACACAGAGCTTGTGGACAAGACGTGGAGGGCCACTAGAGCCGCCCTACCC	60			
Db	215	CATGTCGGGGCTCCACACAGAGCTTGTGGACAAGACGTGGAGGGCCACTAGAGCCGCCCTACCC	215			
Qy	61	GCCCCAAGAGCTTTCTTACCCAGTGCAAGATGCCCGGACGGACACAGAGCTGGGGCTTCTT	120			
Db	216	GCCCCAAGAGCTTTCTTACCCAGTGCAAGATGCCCGGACGGACACAGAGCTGGGGCTTCTT	275			
Qy	121	GGAGTACACAGACACACATCCCGGACCTATGCTCCCACTGTGCGCGGGCTCCATCATCCCA	180			
Db	276	GGAGTACACAGACACACATCCCGGACCTATGCTCCCACTGTGCGCGGGCTCCATCATCCCA	335			
Qy	181	GCCCCAGGGGGAGAGGAGCCCTCCCTGCTGTGAATTTCAAGCCCGGAGATGAACAGTCCCA	240			
Db	336	GCCCCAGGGGGAGAGGAGCCCTCCCTGCTGTGAATTTCAAGCCCGGAGATGAACAGTCCCA	395			
Qy	241	GGAAGCTCCAGCTGGGAGCCAGAGTCCCACTACCTGSCCGAGCTGGGGAGATCAGAGAT	300			
Db	396	GGAAGCTCCAGCTGGGAGCCAGAGTCCCACTACCTGSCCGAGCTGGGGAGATCAGAGAT	455			
Qy	301	GGAGTTTCAATTGAAGCAAGAGGCGCTCCGGCTTAGAGCTCTGTGTCACCCCTGCTGGACCC	360			
Db	456	GGAGTTTCAATTGAAGCAAGAGGCGCTCCGGCTTAGAGCTCTGTGTCACCCCTGCTGGACCC	515			
Qy	361	GTCAACCCCTGTGGACACAGGGCCAGAGCTGGGGAGATCTGAAGACTCTCAACCAAGACCGTAG	420			
Db	516	GTCAACCCCTGTGGACACAGGGCCAGAGCTGGGGAGATCTGAAGACTCTCAACCAAGACCGTAG	575			
Qy	421	CTTGACGGGACAGGTGGAACCGGTGTCTCCGCCAGGCCGCCCGGACACTGACCTCTGAGCT	480			
Db	576	CTTGACGGGACAGGTGGAACCGGTGTCTCCGCCAGGCCGCCCGGACACTGACCTCTGAGCT	635			
Qy	481	GGAGCTGTGTCCGCTCACAGGCTGTCTCAAGAGAGAGCTATCCAGAACATGACCGCGTGA	540			
Db	636	GGAGCTGTGTCCGCTCACAGGCTGTCTCAAGAGAGAGCTATCCAGAACATGACCGCGTGA	695			
Qy	541	CCGAGAGATACCAATGGTAAAGACAGAGATCTCTAAGCTAAAGAAAGACAGCAACAGCT	600			
Db	696	CCGAGAGATACCAATGGTAAAGACAGAGATCTCTAAGCTAAAGAAAGACAGCAACAGCT	755			
Qy	601	GGAAGAGAGAGCTGACCAAGCCGAGCCCGGAGCTGAGAAACCGGTGTCAACCGCGCCCATCGA	660			
Db	756	GGAAGAGAGAGCTGACCAAGCCGAGCCCGGAGCTGAGAAACCGGTGTCAACCGCGCCCATCGA	815			
Qy	661	GTCGAGACACCGACAGCTGTGTGACATCATCTACGACGAGAACCGGAAAGAGCTGAAGC	720			
Db	816	GTCGAGACACACCGACAGCTGTGTGACATCATCTACGACGAGAACCGGAAAGAGCTGAAGC	875			
Qy	721	TGCAATTCGGAATTCGGAAGGCGCTGGGGGCCCAAGGTGGAAGCTGCGGTGACCAACAGACC	780			
Db	876	TGCAATTCGGAATTCGGAAGGCGCTGGGGGCCCAAGGTGGAAGCTGCGGTGACCAACAGACC	935			

QY 781 CTCGACACCCGCGAGTATCATGAGAACATCAAAATAAACGAGCGATGCGGAGAGAGCT 840  
Db 996 CTCGACACCCGCGAGTATCATGAGAACATCAAAATAAACGAGCGATGCGGAGAGAGCT 995  
QY 841 AATCTTTATCTTCAAGAGAGAGAAATCACGCTGCGAAACAAATGGAACGAAATTTCTGCA 900  
Db 996 AATCTTTATCTTCAAGAGAGAGAAATCACGCTGCGAAACAAATGGAACGAAATTTCTGCA 1055  
QY 901 GCGCTATGACACGACTCATATGAGAGGCTTTGGAAAAAAAAGTGAAGCCATCGAAAAACACC 960  
Db 1056 GCGCTATGACACGACTCATATGAGAGGCTTTGGAAAAAAAAGTGAAGCCATCGAAAAACACC 1115  
QY 961 GCGCGCGCGCGCCAAAGAGAGCAAGGTGCGAGTACTAGAAAACAGATTCTTGAGAT 1020  
Db 1116 CCGGCGCGCGCGCCAAAGAGAGCAAGGTTCGAGTACTAGAAAACAGATTCTTGAGAT 1175  
QY 1021 CCGCAAGCAGCCCGAGCTGACAGAGCGCATGCAAGCAGGCTGGGCGAGCGGGCGAGTGG 1080  
Db 1176 CCGCAAGCAGCCCGAGCTGACAGAGCGCATGCAAG--AGGTTGGGCGAGCGGGCGAGTGG 1232  
QY 1081 GCTGTTCATGTGCGGCGCGCCGAGCAGACAGAGGTGTCAGAGATCATGTGCGCTCTC 1140  
Db 1233 GCTGTTCATGTGCGGCGCGCCGAGCAGACAGAGGTGTCAGAGATCATGTGCGCTCTC 1292  
QY 1141 AGAGCAGAGAACTTGAAGACAGATGCGCAGCTGCGCTGATCCGCGCATGCTGTA 1200  
Db 1293 AGAGCAGAGAACTTGAAGACAGATGCGCAGCTGCGCTGATCCGCGCATGCTGTA 1352  
QY 1201 CGAAGCTGACCAAGCCGATCAAGTTTCATCAACATGAAACGGGCTTATGCGCGACCTT 1260  
Db 1353 CGAAGCTGACCAAGCCGATCAAGTTTCATCAACATGAAACGGGCTTATGCGCGACCTT 1412  
QY 1261 GAAAGTGTACAAAGACCCGCGAGTCAATGAAATGTGAGTGAAGAGAAAGAACTTT 1320  
Db 1413 GAAAGTGTACAAAGACCCGCGAGTCAATGAAATGTGAGTGAAGAGAAAGAACTTT 1472  
QY 1321 CCGGAGAAAGTTTATGACAGCATCCCAAGAACTTTGGCTGATCGATCTTCTGAGAG 1380  
Db 1473 CCGGAGAAAGTTTATGACAGCATCCCAAGAACTTTGGCTGATCGATCTTCTGAGAG 1532  
QY 1381 GAAAGCAGTGGCTGAGTGGTCTCTATTACTTACCTGACTAAGAAAGAACTTATTA 1440  
Db 1533 GAAAGCAGTGGCTGAGTGGTCTCTATTACTTACCTGACTAAGAAAGAACTTATTA 1592  
QY 1441 GAGCTGTGAGACGAGAGCTATCGGCGCCGCGCAAGAGCCAGCAGCAACAAGACGA 1500  
Db 1593 GAGCTGTGAGACGAGAGCTATCGGCGCCGCGCAAGAGCCAGCAGCAACAAGACGA 1652  
QY 1501 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560  
Db 1653 GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1712  
QY 1561 AGATGAGAAAGAGAAAGAAAGAGCGGAGAAAGAGAGAGAGAGAGAGAGAGAGAG 1620  
Db 1713 AGATGAGAAAGAGAAAGAAAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1772  
QY 1621 CGACAGAGAAAGCTTCTCAAGAGAAAGCAGACGACCTCAGGAGAGAGAGAGAGAG 1680  
Db 1773 CGACAGAGAGAGCTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1832  
QY 1681 GAAAGAGCTGTGCGCTCCAAAGAGCGCAAACTGCGCAACAGCCAGAGAGAGAGAG 1740  
Db 1833 GAAAGAGCTGTGCGCTCCAAAGAGCGCAAACTGCGCAACAGCCAGAGAGAGAGAG 1892  
QY 1741 CCGCATCAACCCGCTCAATGGCTTAATGAGGCCCAACAGCCAGAGAGAGAGAGAG 1800  
Db 1893 CCGCATCAACCCGCTCAATGGCTTAATGAGGCCCAACAGCCAGAGAGAGAGAGAG 1952  
QY 1801 GAGCGCCAGCTGCGCTCCATGAGAGCTGAATGAGATTTCTGCTGAGAGAGAGAGAG 1860  
Db 1953 GAGCGCCAGCTGCGCTCCATGAGAGCTGAATGAGATTTCTGCTGAGAGAGAGAGAG 2012  
QY 1861 GGAAGACGCAAGAAAGATCTCTGAGAACAGGCGCGCAACTGATCGCCAGT 1920

Db 2013 GGAAGACGCAAGAAAGATCTCTCGAACAAGCGCGCAATGCTGCGCATGCGCCGAT 2072  
QY 1921 GGTGGGCTCCAAAGCTGTGTGCGAGTGTAAAGACTTCTACTTCAACTCAAGAAAGGCA 1980  
Db 2073 GGTGGGCTCCAAAGCTGTGTGCGAGTGTAAAGACTTCTACTTCAACTCAAGAAAGGCA 2132  
QY 1981 GAACTCGATGATGATCTTTCAGCAGCAGCAAGCTGAGATGAGAGAGAGAGAGAGAG 2040  
Db 2133 GAACTCGATGATGATCTTTCAGCAGCAGCAAGCTGAGATGAGAGAGAGAGAGAGAG 2192  
QY 2041 GAGAGAAAGAAAG 2100  
Db 2193 GAGAGAAAGAAAG 2252  
QY 2101 GAGAGATGAGAGATGAG 2160  
Db 2253 GAGAGATGAGAGATGAG 2312  
QY 2161 GAGTGAAGCTTTTACATGCGCTCTGAGAAATGAGGTGCCAGAGAGAGATGAGTGGCCAGC 2220  
Db 2313 GAGTGAAGCTTTTACATGCGCTCTGAGAAATGAGGTGCCAGAGAGAGATGAGTGGCCAGC 2322  
QY 2221 CACTGTCAACAAAGCTGAGACACGAGAGATGCCCTCTCTCTCAACTGAGAGCGCGCA 2280  
Db 2323 -ACTGTCAACAAAGCTGAGACACGAGAGATGCCCTCTCTCTCAACTGAGAGCGCGCA 2381  
QY 2281 GACACAGGAGAGAAATGGGCGCCAGAGCCCGCAGCACCCTGAGGCGCGAGCGGAGCAGCC 2340  
Db 2382 GACACAGGAGAGAAATGGGCGCCAGAGCCCGCAGCACCCTGAGGCGCGAGCGGAGCAGCC 2441  
QY 2341 AAGGCCACCCACCCACAGAGAGACATCCGAGGCGCCCATTTGAGCCACCCGCGCTC 2400  
Db 2442 AAGGCCACCCACCCACAGAGAGACATCCGAGGCGCCCATTTGAGCTCACCCGCGCTC 2501  
QY 2401 TGAAGCCACCGAGCGCCCTTACGCCCCCAGCAGCACCCTCTCTGCACTGAGCGCTCC 2460  
Db 2502 TGAAGCCACCTTACGCGCTTACGCCCCCAGCAGCACCCTCTCTTTCATCTCTCC 2561  
QY 2461 TGTGTCTCCCAAG 2520  
Db 2562 TGTGTCTCCCAAG 2621  
QY 2521 GAG 2580  
Db 2622 GAG 2681  
QY 2581 GCCGCTCAAGAGCTGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
Db 2682 GCCGCTCAAGAGCTGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2741  
QY 2641 GAGAGCGCTGAG 2700  
Db 2742 GAGAGCGCTGAG 2801  
QY 2701 CGGCAAGGCGCAACTGCGCAAGAGAGTCTGAGGCGCCCGCAGAGAGAGAGAGAGAG 2760  
Db 2802 CGGCAAGGCGCGCAACTGCGCAAGAGAGTCTGAGGCGCCCGCAGAGAGAGAGAGAGAG 2861  
QY 2761 CTGCAATGCAAG 2820  
Db 2862 CTGCAATGCAAG 2921  
QY 2821 AAGGCCAGAGCTCTTCAACCCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
Db 2922 AAGGCCAGAGCTCTTCAACCCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2981  
QY 2881 ACTGAGCTTGAAGCAGTGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940  
Db 2982 ACTGAGCTTGAAGCAGTGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3041  
QY 2941 AGTCCATAG 3000

Db 3042 AGTCATGACCCCCCGGAGGACGACGCTCCACCAACGACTCCGCCACC 3101  
Qy 3001 GCCACGGCAAAACCTGACGCGGAGAGCGACGCGCCCTCAGACGCTGGCAGACGCCCG 3060  
Db 3102 GCCACGGCAAAACCTGACGCGGAGAGCGACGCGCCCTCAGACGCTGGCAGACGCCCG 3161  
Qy 3061 GGGCAAGAGCAGGAGCCCGGACACCCCGCGGCAAGAGAG----- 3100  
Db 3162 GGGCAAGAGCAGGAGCCCGGACACCCCGCGGCAAGAGAGCAGTGTGTTCTT 3221  
Qy 3101 ----GCTTTCGACGCGGAGCCCAAGAGCTGCTGGGAGACCCCTTGTGACCTTCGG 3156  
Db 3222 CCCAGGCTTTCGACGCGGAGCCCAAGAGCTGCTGGGAGACCCCTTGTGACCTTCGG 3281  
Qy 3157 CCGGCTTTCGCGGAGCGCGCGCGGAGGATCAAGAGCTTCCCGGATGCGCGAGCC 3216  
Db 3282 CCGGCTTTCGCGGAGCGCGCGCGGAGGATCAAGAGCTTCCCGGATGCGCGAGCC 3341  
Qy 3217 CTCAGGCTTTCCTACGCTCACCCTGGTCAACCACTGCGCCCTGGGCTCCATGACACTGC 3276  
Db 3342 CTCAGGCTTTCCTACGCTCACCCTGGTCAACCACTGCGCCCTGGGCTCCATGACACTGC 3401  
Qy 3277 CCGGCGGCTTCGCGCGGCGCCACCACTTCCTCAACCGGCTTCCTCATCTCTGCG 3336  
Db 3402 CCGGCGGCTTCGCGGCGCCACCACTTCCTCAACCGGCTTCCTCATCTCTGCG 3461  
Qy 3337 CAAGCACCCCAAGGCTCTGAGAGGCAAAAGTGCCATCTCCCAAGGAATGTGCGTCA 3396  
Db 3462 CAAGCACCCCAAGGCTCTGAGAGGCAAAAGTGCCATCTCCCAAGGAATGTGCGTCA 3521  
Qy 3397 GCTCCAGTCCCGTACTCAGAGCATGCAAGGCGCGCGTGGGCGCTGTCACTAGGGGCT 3456  
Db 3522 GCTCCAGTCCCGTACTCAGAGCATGCAAGGCGCGCGTGGGCGCTGTCACTAGGGGCT 3581  
Qy 3457 GCCCGTCCCATGAGACCCCAAAAGCTGGGACCTTCAGCGAGTGAAGCAGGACGCT 3516  
Db 3582 GCCCGTCCCATGAGACCCCAAAAGCTGGGACCTTCAGCGAGTGAAGCAGGACGCT 3641  
Qy 3517 GTTCCCAACGGGAGCAGGCTGGGCGCACCGAGAGCTGGGGGTGCCCAACGCCAGAGGCG 3576  
Db 3642 GTTCCCAACGGGAGCAGGCTGGGCGCACCGAGAGCTGGGGGTGCCCAACGCCAGAGGCG 3701  
Qy 3577 GTCCGTGCTGAGAGGAGCAGCTCTGGGCTCAGTTCCGGGCGGAGCATCAACAAAGCAT 3636  
Db 3702 GTCCGTGCTGAGAGGAGCAGCTCTGGGCTCAGTTCCGGGCGGAGCATCAACAAAGCAT 3761  
Qy 3637 TCCCAAGCACAGGAGTGCCTTCGACAGCGGCATCACTACCGGCTCATCAACCGG 3696  
Db 3762 TCCCAAGCACAGGAGTGCCTTCGACAGCGGCATCACTACCGGCTCATCAACCGG 3821  
Qy 3697 CACGCCAGCTGACGTCTGTATCAAGGGCACCATCAACAGATCATCGGCAAGAGCAGCCC 3756  
Db 3822 CACGCCAGCTGACGTCTGTATCAAGGGCACCATCAACAGATCATCGGCAAGAGCAGCCC 3881  
Qy 3757 GAGTCCGTTGAGACGCGGAGCGGAGGAGCAGCTGCGCCAAAGGCGACGTCATCTTACAAGG 3816  
Db 3882 GAGTCCGTTGAGACGCGGAGCGGAGGAGCAGCTGCGCCAAAGGCGACGTCATCTTACAAGG 3941  
Qy 3817 CAAGAAAGGCGCAGCTTGTCTATGAGGGTGAATGTCTGTGACCCAGTGTCCAAGA 3876  
Db 3942 CAAGAAAGGCGCAGCTTGTCTATGAGGGTGAATGTCTGTGACCCAGTGTCCAAGA 4001  
Qy 3877 GGAACGGCAGAAAGCTCAGAGACCCCCCATGAGAGCGGCGCGCCCAAGGCGACCTATGA 3936  
Db 4002 GGAACGGCAGAAAGCTCAGAGACCCCCCATGAGAGCGGCGCGCCCAAGGCGACCTATGA 4061  
Qy 3937 CATGATGAGAGGCGCGGTGGGAGAGCAGCTCTCAGCAGAGTCAGAGTCTCATGGG 3996  
Db 4062 CATGATGAGAGGCGCGGTGGGAGAGCAGCTCTCAGCAGAGTCAGAGTCTCATGGG 4121  
Qy 3997 CCGTGCATTCGCGCGGAGGAGCAGACGCGCCCACTTCAAAAGAGCAGCAGCATCG 4056  
Db 4122 CCGTGCATTCGCGCGGAGGAGCAGACGCGCCCACTTCAAAAGAGCAGCAGCATCG 4181

Qy 4057 CCGGTCCATCAACAAGGATTCCTCGGTCTACGTGAGGACAGAGAGCATACCTGCG 4116  
Db 4182 CCGGTCCATCAACAAGGATTCCTCGGTCTACGTGAGGACAGAGAGCATACCTGCG 4241  
Qy 4117 TCGGAGGCCAAGCTCTTAAAGCGGAGGAGCAGCGCTCGCGCCCAACGCGCTTCAACGGGA 4176  
Db 4242 TCGGAGGCCAAGCTCTTAAAGCGGAGGAGGAGCAGCGCTCGCGCCCAACGCGCTTCAACGGGA 4301  
Qy 4177 CTGACCCGAGGCTTACAAAGCAGAGCCCTGGGCGCCCTGGAAGCTGAAGCGCGCCCTGA 4236  
Db 4302 CTGACCCGAGGCTTACAAAGCAGAGCCCTGGGCGCCCTGGAAGCTGAAGCGCGCCCTGA 4361  
Qy 4237 GGGCTGTGTGGCCACGGTGAAGAGGAGGAGCGCTCATCTCATGAGATCCCGGCGAGGA 4296  
Db 4362 GGGCTGTGTGGCCACGGTGAAGAGGAGGAGGAGCGCTCATCTCATGAGATCCCGGCGAGGA 4421  
Qy 4297 GCTGCGGACACGCGCGGAGCTGCGCCCTGGGCGCGGCTCAAGAGAGGCTTCATCAC 4356  
Db 4422 GCTGCGGACACGCGCGGAGCTGCGCCCTGGGCGCGGCTCAAGAGAGGCTTCATCAC 4481  
Qy 4357 GCAAGGACACCCCGCTCAAGTACGACACCGCGCGCTCACTGCTCCAAAAGACGGA 4416  
Db 4482 GCAAGGACACCCCGCTCAAGTACGACACCGCGCGCTCACTGCTCCAAAAGACGGA 4541  
Qy 4417 CGTACGCTCCCTCATGCGGACGCGCGGAGCGTTCCCAACCGGTGACACCGCTGAGTGT 4476  
Db 4542 CGTACGCTCCCTCATGCGGACGCGCGGAGCGTTCCCAACCGGTGACACCGCTGAGTGT 4601  
Qy 4477 GATGCGGACGCGCGGACCTGGAACGTGCTGCTACGAGAGAGCTGGAAGAGCGCGGC 4536  
Db 4602 GATGCGGACGCGCGGACCTGGAACGTGCTGCTACGAGAGAGCTGGAAGAGCGCGGC 4661  
Qy 4537 AGGGAACCGGACGAGCTCGGAGGAGCTTCATTTGCGCGCGCGCGCTCATTTGCTGTA 4596  
Db 4662 AGGGAACCGGACGAGCTCGGAGGAGCTTCATTTGCGCGCGCGCGCTCATTTGCTGTA 4721  
Qy 4597 GCTGGGTAAAGCGCGGAGCGCGCGCTGACCTTATGAGAACAGAGGAGACCTTTGCGCG 4656  
Db 4722 GCTGGGTAAAGCGCGGAGCGCGCGCTTATGAGAACAGAGGAGACCTTTGCGCG 4781  
Qy 4657 CCACTTCCACAGAGTTCGCGCTGTGACATGCGGAGGCCACGCGCGCTGAGAGAGG 4716  
Db 4782 CCACTTCCACAGAGTTCGCGCTGTGACATGCGGAGGCCACGCGCGCTGAGAGAGG 4841  
Qy 4717 CAGCTTTTGTCCAGAGGAGCATCCAGAGCGGAAGCTGACGTGACGCTTGTGAGAT 4776  
Db 4842 CAGCTTTTGTCCAGAGGAGCATCCAGAGCGGAAGCTGACGTGACGCTTGTGAGAT 4901  
Qy 4777 CCGCAAGTCCCGGACAGACCGGTGCGGAGCAGACCCCAACCCCATCTGCGCTTATGA 4836  
Db 4902 CCGCAAGTCCCGGACAGACCGGTGCGGAGCAGACCCCAACCCCATCTGCGCTTATGA 4961  
Qy 4837 GCACTGTCTTCGGGAGGTGAGTGGCTGTGACCTGTATTCGAGGCAGATCCCGTGGCTT 4896  
Db 4962 GCACTGTCTTCGGGAGGTGAGTGGCTGTGACCTGTATTCGAGGCAGATCCCGTGGCTT 5021  
Qy 4897 CGAACCCCACTTCATACCCCGGCGATTCCTTGTGACGACGCGCTGCTTACTGCGC 4956  
Db 5022 CGAACCCCACTTCATACCCCGGCGATTCCTTGTGACGACGCGCTGCTTACTGCGC 5081  
Qy 4957 CCGAACCTGGGCGCGCAACCCCACTTACCGGACCTGTATTCGAGGCAGATCCCGTGGCTT 5016  
Db 5082 CCGAACCTGGGCGCGCAACCCCACTTACCGGACCTGTATTCGAGGCAGATCCCGTGGCTT 5141  
Qy 5017 CTACCCCGACACGCGCGGCTGTGAGAACCGGACGACCATCATATGACTATCATCTTC 5076  
Db 5142 CTACCCCGACACGCGCGGCTGTGAGAACCGGACGACCATCATATGACTATCATCTTC 5201  
Qy 5077 GCAAGCAGATGACACAAACAGGCGCACCGCATGCGCCAGCGAGCTGATATGTGTGAGGG 5136  
Db 5202 GCAAGCAGATGACACAAACAGGCGCACCGCATGCGCCAGCGAGCTGATATGTGTGAGGG 5261



QY 5137 CCTCTGCCCCGAGTCTCGTGGCACTAAGCTGCGGTCCCCGAGGACATCAT 5196  
Db 5262 CCTCTGCCCCCGAGTCTCGTGGCACTCAACTACGCTGCGGGTCCCCGAGGACATCAT 5321  
QY 5197 CGACCTGTCCCAAGGCGCACACTGCTGTGTGCTCGGCCCCCGGACACAGGAGACCCGACG 5256  
Db 5322 CGACCTGTCCCAAGGCGCACACTGCTGTGTGCTCGGCCCCCGGACACAGGAGACCCGACG 5381  
QY 5257 CACCGGCATGAGACCGGCTTGCTTACCTCCCAACCGGCCCCGAGCCCTTGAGAGCGGCA 5316  
Db 5382 CACCGGCATGAGACCGGCTTGCTTACCTCCCAACCGGCCCCGAGCCCTTGAGAGCGGCA 5441  
QY 5317 CAGCAGCTCCCACTCTCTCCCAAGAGTCCAAACACTTGACAAAACAACAACAAGCTG 5376  
Db 5442 CAGCAGCTCCCACTCTCTCCCAAGAGTCCAAACACTTGACAAAACAACAACAAGCTG 5501  
QY 5377 CTGCTCCAGCGGGGACGAGACCGGGATGAGAGCGGGACCGGGATCGGGAGACGGGAAA 5436  
Db 5502 CTGCTCCAGCGGGGACGAGACCGGGATGAGAGCGGGACCGGGATCGGGAGACGGGAAA 5561  
QY 5437 GTTCATCTCAGGTCCACACGACGAGTGGAGACGCAACCATCTGGAGACTGTACAGA 5496  
Db 5562 GTTCATCTCAGGTCCACACGACGAGTGGAGACGCAACCATCTGGAGACTGTACAGA 5621  
QY 5497 GCAGAGCAGCGGACAGCGGCGACAGCGCGGGGTGGGGCGACAGACGCGCGCGC 5556  
Db 5622 GCAGAGCAGCGGACAGCGGCGACAGCGCGGGGTGGGGCGACAGACGCGCGCGC 5681  
QY 5557 CTCCCACTCCCATGCGCACACGACCTGCGCCATCTCCCTCGGACCCAGAGAGCCCTCA 5616  
Db 5682 CTCCCACTCCCATGCGCACACGACCTGCGCCATCTCCCTCGGACCCAGAGAGCCCTCA 5741  
QY 5617 GCAGAGACCCAGTGTGCTTCAACAACAGGAGTGAAGGGTATCATCAACGCTGTGGAGC 5676  
Db 5742 GCAGAGACCCAGTGTGCTTCAACAACAGGAGTGAAGGGTATCATCAACGCTGTGGAGC 5801  
QY 5677 CAGCAAGCCCAAGTGTGAGGTCCACTCCACTCTCTCAACCCGTTGCGCCAGTGGCAC 5736  
Db 5802 CAGCAAGCCCAAGTGTGAGGTCCACTCCACTCTCTCAACCCGTTGCGCCAGTGGCAC 5861  
QY 5737 ATTCCACCTGCGACCCACTGCGCACCTGGGCGGCAACCCCTGATGGGTGTACCTTACCT 5796  
Db 5862 ATTCCACCTGCGACCCACTGCGCACCTGGGCGGCAACCCCTGATGGGTGTACCTTACCT 5921  
QY 5797 CATTGAGCCCGTCTTGTGCTGCCAAGAGGCCCCCGGAGTGGCCCGGACAGAGCGGCCG 5856  
Db 5922 CATTGAGCCCGTCTTGTGCTGCCAAGAGGCCCCCGGAGTGGCCCGGACAGAGCGGCCG 5981  
QY 5857 AGCAGACACCGGCAATGCTTCTTGCGCAAGCCCCCAAGCCGCTCGGGCTGAGGCCG 5916  
Db 5982 AGCAGACACCGGCAATGCTTCTTGCGCAAGCCCCCAAGCCGCTCGGGCTGAGGCCG 6041  
QY 5917 CTGCTCCCGCAGAGAGGTGCGAGCGCGCGGCCCTTAAGTGCCTCTGTGTCTGGGCAAGC 5976  
Db 6042 CTGCTCCCGCAGAGAGGTGCGAGCGCGGCCCTTAAGTGCCTCTGTGTCTGGGCAAGC 6101  
QY 5977 CACCAATGCGCGCACCCCTGCGAAGAACTCGCACTTCAACAAGCCAGGCCCCGAC 6036  
Db 6102 CACCAATGCGCGCACCCCTGCGAAGAACTCGCACTTCAACAAGCCAGGCCCCGAC 6161  
QY 6037 GAGCGCACCTGCTGCTGAGCTCGGACCCGACCGGAAAAAGACTCAAAAGTAAACCTTTTC 6096  
Db 6162 GAGCGCACCTGCTGCTGAGCTCGGACCCGACCGGAAAAAGACTCAAAAGTAAACCTTTTC 6221  
QY 6097 CATTCAGGAATCTGGAATCTCGTTCTCTGGGTTTACCAACGCGACAGTTACAGCCCGGAAG 6156  
Db 6222 CATTCAGGAATCTGGAATCTCGTTCTCTGGGTTTACCAACGCGACAGTTACAGCCCGGAAG 6281  
QY 6157 GGTGAGCCCGTCAAGCCCTGTGAGCTCAACCAAGTCTGACCAAGCAAGAGGGTCCCCAA 6216  
Db 6282 GGTGAGCCCGTCAAGCCCTGTGAGCTCAACCAAGTCTGACCAAGCAAGAGGGTCCCCAA 6341  
QY 6217 GCACCTGGAAGCTGACAAAGGCCACTTGAGAGGGGAGCTGCGGCCCAAGCCAG 6276

Db 6342 GCACCTGGAAGAGCTGACAAAGGCCACTGGAGGGGAGGCTGCCGCCAAGCCAG 6401  
QY 6277 CCGCGTGAAGCTTGGCGGGGAGCGGCCCACTCCCAACACTTGCGGGCGCTGTGAGAG 6336  
Db 6402 CCGCGTGAAGCTTGGCGGGGAGCGGCCCACTCCCAACACTTGCGGGCGCTGTGAGAG 6461  
QY 6337 CCAAGCCTGTGCAGCCCGTGTCTCAGAACGCGCCAGGGGTCAAAAGTCAACAGCGGT 6396  
Db 6462 CCAAGCCTGTGCAGCCCGTGTCTCAGAACGCGCCAGGGGTCAAAAGTCAACAGCGGT 6521  
QY 6397 GGTCAACCTTGGGCCAGCAGTATGAGAGTATCAACAGAGCTTACACCCGGACACCC 6456  
Db 6522 GGTCAACCTTGGGCCAGCAGTATGAGAGTATCAACAGAGCTTACACACCCGGACACCC 6581  
QY 6457 ACAGAGCTCAGCGGACCCCTGCGCGCCCTTACTCTTCCCTGGGGGCGAGCTGCC 6516  
Db 6582 ACAGAGCTCAGCGGACCCCTGCGCGCCCTTACTCTTCCCTGGGGGCGAGCTGCC 6641  
QY 6517 GGTCTGGAACCTCGCGGCCCAACCAAGTGAACCTTACCTCCGCCCCGGAACATGTGC 6576  
Db 6642 GGTCTGGAACCTCGCGGCCCAACCAAGTGAACCTTACCTCCGCCCCGGAACATGTGC 6701  
QY 6577 CCGGCGCGTGGCTCCCCCAGAGGAGGGGCAAGAGTCTCCAGAGCCAAACAAGAC 6636  
Db 6702 CCGGCGCGTGGCTCCCCCAGAGGAGGGGCAAGAGTCTCCAGAGCCAAACAAGAC 6761  
QY 6637 GTGCGTCTTGGGTGGTGGAGAGCGTATTGAACCTGTGTCCTCCACCGGAGGGCATGAC 6696  
Db 6762 GTGCGTCTTGGGTGGTGGAGAGCGTATTGAACCTGTGTCCTCCACCGGAGGGCATGAC 6821  
QY 6697 GAGGCCAGGGCACTCCGGAAGTGTGTGACCCGCTGTGTACCCGAGTGGGAAACAAGC 6756  
Db 6822 GAGGCCAGGGCACTCCGGAAGTGTGTGACCCGCTGTGTACCCGAGTGGGAAACAAGC 6881  
QY 6757 GAGGCCAGCGAGATGGGTCTCAAGTCTTCAGGCAACACAGCCAGCGGCCAGCTTCTT 6816  
Db 6882 GAGGCCAGCGAGATGGGTCTCAAGTCTTCAGGCAACACAGCCAGCGGCCAGCTTCTT 6941  
QY 6817 CAGCAAGCTGACCCGAGAGCAACTCGGCATGAGTCAAGTCCAAAGACAAAGATCAACA 6876  
Db 6942 CAGCAAGCTGACCCGAGAGCAACTCGGCATGAGTCAAGTCCAAAGACAAAGATCAACA 7001  
QY 6877 GAAAGTGAACAACCAACCGGAATGAGCTGAATATACAATATCAGCCAGCTGGAGCGA 6936  
Db 7002 GAAAGTGAACAACCAACCGGAATGAGCTGAATATACAATATCAGCCAGCTGGAGCGA 7061  
QY 6937 GATCTTCAATATGCCCCGATTCACCGGAACAGGCTTATGACTTATAGAGCAGGCGGT 6996  
Db 7062 GATCTTCAATATGCCCCGATTCACCGGAACAGGCTTATGACTTATAGAGCAGGCGGT 7121  
QY 6997 GCAGGAACATGCGACACCAACATGGGGCTGGAGGCAATTAAGAAAGGCACTATGGG 7056  
Db 7122 GCAGGAACATGCGACACCAACATGGGGCTGGAGGCAATTAAGAAAGGCACTATG-- 7179  
QY 7057 TAAATATGACCAAGTGGGAAGTCCCCGCGCTCAGCGCCAATGCTTTTAAACCTCTGAA 7116  
Db 7180 ----- 7179  
QY 7117 TGCCAGTGCAGGCTGCGCGCTGTATGCAATTAACGCTGTGTGACGAGAGTGAACA 7176  
Db 7180 ----- 7179  
QY 7177 CACACTCACTCGCAGGATGGGAGGGAAGGCCAAGGCTCTGGCAGAACCAAGAGCG 7236  
Db 7180 -----GAGTGGAGGGAAGGCCAAGGCTCTGGCAGAACCAAGAGCGG 7223  
QY 7237 AAAAGCCAAAGTCCCCGAGCCCGGAGCTTGCAATCTGGGAGCCGAGCAACCTGTGTCTCTC 7296  
Db 7224 AAAAGCCAAAGTCCCCGAGCCCGGAGCTTGCAATCTGGGAGCCGAGCAACCTGTGTCTCTC 7283  
QY 7297 AGTGCACCTCGAGGAGACTGCAACCGGCGGACGCGCTCACCAACCGGCTGTGGAGGA 7356



Db	7284	AGTGCATCTCGAAGGGAGACTGCAACCGCCGAGCCGCTCACCAACCGCGTGTGGAGGA	7343
Oy	7357	CAGGCGCTCTGTCCGAGGTTCCACGCGCATTTCCCTTACAACCCCTCGATCATGCGGCTGCA	7416
Db	7344	CAGGCGCTCTGTCCGAGGTTCCACGCGCATTTCCCTTACAACCCCTCGATCATGCGGCTGCA	7403
Oy	7417	GGCGGGTGTCAATGAGTTTCCCAACCCCAACGGGAGTCTCCCGCGGGCAACGGGAGCCCTTGGC	7476
Db	7404	GGCGGGTGTCAATGAGTTTCCCAACCCCAACGGGAGTCTCCCGCGGGCAACGGGAGCCCTTGGC	7463
Oy	7477	TGGCCCCCAACACCGCTGGGAGCAAGAGACCCAAAGCATGTCCTGTGCTGCAGTACAGAC	7536
Db	7464	TGGCCCCCAACACCGCTGGGAGCAAGAGACCCAAAGCATGTCCTGTGCTGCAGTACAGAC	7523
Oy	7537	ACTCTCCGACACGAGTGACTCAGAACAGGGCGGGGGGGGGCGGGCGGTBUBERSPERLUG	7596
Db	7524	ACTCTCCGACACGAGTGACTCAGAACAGGGGGGGGGGGGGCGGGCGG-----TG	7573
Oy	7537	TCAGGTCCCAAGCAACCAAGGAACGGCCCTTGCAGAGACGGGGGGGGCTGCGACTCCCC	7656
Db	7574	TCAGGTCCCAAGCAACCAAGGAACGGCCCTTGCAGAGACGGGGGGGGGGCTGCGACTCCCC	7633
Oy	7657	AACCAAGGAAGAGAGCCCGCTGAAGTCCGAGCCCTTGCATTCATCTGTCCGTCCAGAGCGC	7716
Db	7634	AACCAAGGAAGAGAGCCCGCTGAAGTCCGAGCCCTTGCATTCATCTGTCCGTCCAGAGCGC	7693
Oy	7717	GCATCTCTTGCCTGTCTTAAAGCTTTAACTAAAGCTCCCGCCCGGGCTGGCCCTGTGCAGA	7776
Db	7694	GCATCTCTTGCCTGTGTCTTAAAGCTTTAAAGCTCCCGCCCGGGCTGGCCCTGTGTGCAGA	7753
Oy	7777	CCTTACTCAGGGGAGTGTTTTACCTGTGTCTGCGAAGGGAGGGGAAGGGGCGGAGAGGGG	7836
Db	7754	CCTTACTCAGGGGAGTGTTTTACCTGTGTGTCTGCGAAGGGGAAGGGGCGGAGAGGGG	7813
Oy	7837	GCAGGGGAGGGGTGTGTGGAGGCAACACAAAGCGCGCAGGGGGCCAGGGAAACCAAGACAG	7896
Db	7814	GCAGGGGAGGGGTGTGTGGAGGCAACACAAAGCGCGCAGGGGGCCAGGGAAACCAAGACAG	7873
Oy	7897	GATGACACACGACCTTCCAGCGCACCTGCTCCCGGAAATGACTTTTGGAAACCAAGCTTAAA	7956
Db	7874	GATGACACACGACCTTCCAGCGCACCTGCTCCCGGAAATGACTTTTGGAAACCAAGCTTAAA	7933
Oy	7957	CTGAGCTCGACGCCCCCGGCGCCCTTCCCTCGGCTTCCCATCCCGCTTGAAGGCTTGGACAG	8016
Db	7934	CTGAGCTCGACGCCCCCGGCGCCCTTCCCTCGGCTTCCCATCCCGCTTGAAGGCTTGGACAG	7993
Oy	8017	ATGGAACGAGGCGCCCTGTCTTCAACCCCCCAATGGAGGCTCTTCCCGTCCCCACAGACTGCCCA	8076
Db	7994	ATGGAACGAGGCGCCCTGTCTTCAACCCCCCAATGGAGGCTCTTCCCGTCCCCACAGACTGCCCA	8053
Oy	8077	GCCAAACGAGATTTGCTGAGAAACCAAGTACAGGCGCAGGTGGGCGGACAAAGAGGCGCAGTGGC	8136
Db	8054	GCCAAACGAGATTTGCTGAGAAACCAAGTACAGGCGCAGGTGGGCGGACAAAGAGGCGCAGTGGC	8113
Oy	8137	GCCTGGGGGGAACGGAATGCTCCGAGGACCTGACCTGTTTTCACACATCGTTGGCCGACG	8196
Db	8114	GCCTGGGGGGAACGGAATGCTCCGAGGACCTGACCTGTTTTCACACATCGTTGGCCGACG	8173
Oy	8197	CGGTGGGGAAGAAAGGCGAGATGTAAATGTGTGTTTTCAGAGGTTAATTTTGTATAC	8256
Db	8174	CGGTGGGGAAGAAAGGCGAGATGTAAATGTGTGTTTTCAGAGGTTAATTTTGTATAC	8233
Oy	8257	CTTCAATGAATTAATTAATGATGTTTTCAGCAAGGAAGGACTTACCAGATATTACTCGCGC	8316
Db	8234	CTTCAATGAATTAATTAATGATGTTTTCAGCAAGGAAGGACTTACCAGATATTACTCGCGC	8293
Oy	8317	TGTGCTTTGATCTCTGCTTACCGTTCAAGAGGCGTGTGCAGGCGCGACAGTCCGATGACCC	8376
Db	8294	TGTGCTTTTGTATCTCTGCTTACCGTTCAAGAGGCGGTGTGCAGGCGCGACAGTCCGATGACCC	8353
Oy	8377	CATCATCTGCAGACCAAGGGGGGGGGGACCTGTCTGTACAGCCCGCGTGTCTCTCCCTC	8436
Db	8354	CATCATCTGCAGACCAAGGGGGGGGGGACCTGTCTGTACAGCCCGCGTGTGTCTCTCCCTC	8413

QY	8437	CTCTCCCTTCTTGGCGAGATGAATGAATGCGATTCTGTGGCCGCAATTTGGCGCAGG	8496
Db	8414	CTCTCCCTTCTTGGCGAGATGAATGAATGCGATTCTGTGGCCGCAATTTGGCGCAGG	8473
QY	8497	TGGGTGATTCTGTCATTTACACAGCGTCGTTCTTAATTAAGGAATTAATCTCAAA	8556
Db	8474	TGGGTGATTCTGTCATTTACACAGCGTCGTTCTTAATTAAGGAATTAATCTCAAA	8553
QY	8557	AAAAAAAAAAAAAAAA 8571	
Db	8534	AAAAAAAAAAAAAAAA 8548	
RESULT 5	HSU37146	5989 bp	linear
LOCUS	HSU37146	Human silencing mediator of retinoid and thyroid hormone action	
DEFINITION	(SMART) mRNA, complete cds.		
ACCESSION	U37146		
VERSION	U37146.1	GI:1045654	
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1. (bases 1 to 5989) Chen, J.D. and Evans, R.M.		
TITLE	A transcriptional co-repressor that interacts with nuclear hormone receptors		
JOURNAL	Nature 377 (6548), 454-457 (1995)		
PUBMED	7566127		
REFERENCE	2. (bases 1 to 5989) Chen, J.D. and Evans, R.M.		
AUTHORS	Direct Submission		
JOURNAL	Submitted (27-SEP-1995) J. Don Chen, Gene Expression Lab, The Salk Institute, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA		
FEATURES	location/Qualifiers		
source	1. 5989		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/cell_line="HeLa"		
	1. 5989		
	/gene="SMRT"		
	496.. 4983		
	/gene="SMRT"		
	/note="transcriptional co-repressor"		
	/codon_start=1		
	/product="silencing mediator of retinoid and thyroid hormone action"		
	protein_id="AAC50226.1"		
	/db_xref="GI:1045655"		
	/translation="MEAWMDAPDKAEFAAEAOQLPGDPPCMTSGLPVPVPREVIKAS PHADBSAFSAPBPGHPLGLHDPTARPVLPRTTISNPPLTSSAGHPSVLSROGA ISGMSVOLHVPVSEBAKAVGVGTWGLPLPMOPKCLAPSGVKOEQLSRGQAGPE SLCPVPAOEAASVLRGALSGVGSITKGIISRNVSDAITRGSTTHGTPADVLYK GTTTRTIGBDSRRLDRGRDLSPKGHVITYEGKKHVLSTEGKMSVYQCKEGRSS GPHETVAEPRTYDMMEGRGRAISSASIEGLMGRAIPERHSPHMLKEOHNRGSTT QGIPRSVEAOEDYLRREALTKRGEPPPPSRDLTEAYKVALGLPKAKHGL VAATYKAGRSIHIEPREELHTPELPLAPRLKEGSTTQCTPLKDYDGAATGSKKD VRSIGSGRTPPVPLDVMADAPALERACYESLSKRSKGTSAGSAGSARGAAPVY PELGRKPSPIYEDHGAAPGALPGSGPVYMBEPRLOEGSLSSKASODKLTST PRETIASPHSTVPEHHPHPSPIYEHILRGVSGDYLRSHPPLAPDPSTIRGIPLDA AAATYPLHLAHPNTPYPLVPPYILRGVPTALENNQTINDYITSQMHNTATMAA QORDMIRGLSPRESSLALNAAAPRGIDLSQVPHLVLPPTPGATMDRLAYLP TAQDFMSGRSSSPLSPGCTTHLTKPTTSSSRERDRERDRERERKSLTSTTT VEAAPIWRPGETEOSGSGSSGSSSRPASHSHOHSPISPRDOLLOORPSVL HNGCMGCTIAYVAPSKPTVLRSTSSPVAPPAATPAPATCPICGLTIDGVYPTLMEPV LLEKAPRVARPERPRADTGHAFAPLAPPAASGLEPSSPEKSGEPPPLVPVVGATII LRPAPKANLAAHASPDPAPAPASDPHREKTSKPSIDELRSITGHSSYSPEG VEAPSPVSSSLTHDKGLPKHLBELDKSHLEGLRPQDPPVVLGGSEAAHLPLRLPLP		

ESQSSSSPLLQTAPEVKHQIVVTLAQHISVITODYTRHHPQULSAPLPADLYSPFG  
ASCPVLDIRRPSDLYLPDPDHGAPARSPHEGKRSPEPKPTSVLGGEGEIGIPVS  
PPEGMTEBHSRAVYPLLYRDGEQTEBPMGSKSPGTSOPAFSKLTENSAMVX  
SKJOINKKLNTHNNEBPYNI SOPETL FNNPALTGGLMTRYRQSOA VOEHSTJMG  
EALIRKALMGKDWMEBSPILSANA PNPLINASLI.PAMPITRAGRSHNLTSPGG  
GKAKTSGRSPSSKAKSPAPGLASGRPPSVSVSHSEGDCKNRTPLTNRMWERPSAG  
STPFPYNNPLIMRLQGMVMSPPPGLIPASGPLAHPHMADEPKPLCSQYETLSDS  
E+  
4483. .4620  
/gene="SMRT"  
/note="alternatively spliced insert"

exon  
63.7%; Score 5455; DB 8; Length 5989;  
Query Match 99.7%; Pred. No. 0;  
Best Local Similarity 1; Mismatches 5; Indels 10; Gaps 1;  
Matches 5477; Conservative

QY	3079	GGCACCCCCCGCCGACAAAGAGAGCCCTTGGCAGCCGAGGCCGCAAGAGCTCCCTGGGAGACCC	3138
DB	507	GGACGCCCAACCCCGACAAAGAGAGCCCTTGGCAGCCGAGGCCGCAAGAGCTCCCTGGGAGACCC	566
QY	3139	CCCTTGCTGAGACTTCGCGGCGCTGCCCTTCCCGTGCCCCCGGTGAGGTGATCAAGGCGCTC	3198
DB	567	CCCTTGCTGAGACTTCGCGGCGCTGCCCTTCCCGTGCCCCCGGTGAGGTGATCAAGGCGCTC	626
QY	3199	CCCGATGCCCCCGGACCCCTTCAAGCTTCTCTACGCTCCACCTTGATCAACCACTGCCCCCT	3258
DB	627	CCCGATGCCCCCGGACCCCTTCAAGCTTCTCTACGCTCCACCTTGATCAACCACTGCCCCCT	686
QY	3259	GGGCGCTCCATGACATGCGCGCGCGCGCTCTGCGCGCGCGCCACCAACCATCTCCACCGCGC	3318
DB	687	GGGCGCTCCATGACATGCGCGCGCGCGCTCTGCGCGCGCGCCACCAACCATCTCCACCGCGC	746
QY	3319	TCCCTCATCTCTCTGCGCAAGACACCCAGCGCTCTCGAGAGGCAAAATAGTGCCATCTC	3378
DB	747	TCCCTCATCTCTCTGCGCAAGACACCCAGCGCTCTCTCGAGAGGCAAAATAGTGCCATCTC	806
QY	3379	CCAAGGAATGTGGTCCAGCTCCAGCTCCCGTACTCGAGACATGCCAAAGCCCGGTGGG	3438
DB	807	CCAAGGAATGTGGTCCAGCTCCAGCTCCCGTACTCGAGACATGCCAAAGCCCGGTGGG	866
QY	3439	CCCTGTCAACATGCGGCGCTGCCCTGCCATGGAACCCCAAAAAGCTTGACCCCTTCAAGCGG	3498
DB	867	CCCTGTCAACATGCGGCGCTGCCCTGCCATGGAACCCCAAAAAGCTTGACCCCTTCAAGCGG	926
QY	3499	AGTGAAGCAGAGCAGCTGTCCCAAGCGGCGCAAGGCTGGGCGCACCGAGAGCCTGGGGGT	3558
DB	927	AGTGAAGCAGAGCAGCTGTCCCAAGCGGCGCAAGGCTGGGCGCACCGAGAGCCTGGGGGT	986
QY	3559	GGCCACAGCCGAGAGCGCTCGGTGCTGAGAGGACAGCTCTGGGCTCAGTTCCGGGCGG	3618
DB	987	GGCCACAGCCGAGAGCGCTCGGTGCTGAGAGGACAGCTCTGGGCTCAGTTCCGGGCGG	1046
QY	3619	AAGCATCAACAAAGGCAATTCACAGACACAGGGTGCTCTTGACACAGGCGCATCATACCG	3678
DB	1047	AAGCATCAACAAAGGCAATTCACAGACACAGGGTGCTCTTGACACAGGCGCATCATACCG	1106
QY	3679	CGGCTCCATCAACCGCAGCGCAGCTGACGCTCTGTAACAAGGGACCATCAACAGAGAT	3738
DB	1107	CGGCTCCATCAACCGCAGCGCAGCTGACGCTCTGTAACAAGGGACCATCAACAGAGAT	1166
QY	3739	CATCGGCGAGACAGCCCGAGTCCGCTTGGACCGCGGCGGAGAGACAGCCTGCCAAAGG	3798
DB	1167	CATCGGCGAGACAGCCCGAGTCCGCTTGGACCGCGGCGGAGAGACAGCCTGCCAAAGG	1226
QY	3799	CCAGCTCATCTACGAAGGCAAGAAGGCGCAAGCTTTGTCTTATGAGGGTGCGATGTCTGT	3858
DB	1227	CCAGCTCATCTACGAAGGCAAGAAGGCGCAAGCTTTGTCTTATGAGGGTGCGATGTCTGT	1286
QY	3859	GACCGAGTCTCAAGAGAGAGCGGACGAAGACGCTCAAGACCCGCCCATGAGACGGCGCG	3918
DB	1287	GACCGAGTCTCAAGAGAGAGCGGACGAAGACGCTCAAGACCCGCCCATGAGACGGCGCG	1346

QY	3919	CCCAAGCGCACCTTATGACATGATGAGAGGCGCGCTGGGCGAGAGCCATCTCTCAAGCCAG	3978
DB	1347	CCCAAGCGCACCTTATGACATGATGAGAGGCGCGCTGGGCGAGAGCCATCTCTCTCAAGCCAG	1406
QY	3979	CATGAAGTCTCAAGGCGCGGCGCATCCCGCGGAGGAGACACAGCCGCCACCATCTCA	4038
DB	1407	CATGAAGTCTCAAGGCGCGGCGCATCCCGCGGAGGAGACACAGCCGCCACCATCTCA	1466
QY	4039	AGAGGAGACACACATCCGCGGCTCCATGACACAAAGGATCCCTCGGTCTTACGTGAGAGC	4098
DB	1467	AGAGGAGACACACATCCGCGGCTCCATGACACAAAGGATCCCTCGGTCTTACGTGAGAGC	1526
QY	4099	ACAGAGAGACTACCTGCGTGGAGGCGCAAGCTCTTAAAGCGGAGAGGCGACCTCCGCGC	4158
DB	1527	ACAGAGAGACTACCTGCGTGGAGGCGCAAGCTCTTAAAGCGGAGAGGCGACCTCCGCGC	1586
QY	4159	CCACCGGCTTCAAGGAGCTGACCGAGGCGCTAACAAAGCGAGGCGCTGGGCGCCCTGAA	4218
DB	1587	CCACCGGCTTCAAGGAGCTGACCGAGGCGCTAACAAAGCGAGGCGCTGGGCGCCCTGAA	1646
QY	4219	GCTGAAGCGGCGCATGAGAGGCGTGTGCGCACGCTGTAAGAGAGGCGGCGCTCATCCA	4278
DB	1647	GCTGAAGCGGCGCATGAGAGGCGTGTGCGCACGCTGTAAGAGAGGCGGCGCTCATCCA	1706
QY	4279	TGAGATCCCGCGGAGAGCTTGGGCAACGCTCCGAGGCTGCTGCTGCTGCTTCAAGAGA	4338
DB	1707	TGAGATCCCGCGGAGAGCTTGGGCAACGCTCCGAGGCTGCTGCTGCTGCTTCAAGAGA	1766
QY	4339	CAAGAGGGCTCATACGCGAGGGGACCCCGCTCAAGATACGACACCGGCGGCTCCACAC	4398
DB	1767	CAAGAGGGCTCATACGCGAGGGGACCCCGCTCAAGATACGACACCGGCGGCTCCACAC	1826
QY	4399	TGGCTTCAAAAAGACGACGTACGCTCCCTCATTCGCGAGCCCGGCGGACGTTCCACC	4458
DB	1827	TGGCTTCAAAAAGACGACGTACGCTCCCTCATTCGCGAGCCCGGCGGACGTTCCACC	1886
QY	4459	CGTGACCCCGCTGATGTGATGCGGACGCGCGGCGACTGGAAGTGTCTGCTTCAAGAGA	4518
DB	1887	CGTGACCCCGCTGATGTGATGCGGACGCGCGGCGACTGGAAGTGTCTGCTTCAAGAGA	1946
QY	4519	GAGCGTGAAGAGCGGCGGACGCGGACGCGTGGGGGGCTCATTTGGCGGGGGCG	4578
DB	1947	GAGCGTGAAGAGCGGCGGACGCGGACGCGTGGGGGGCTCATTTGGCGGGGGCG	2006
QY	4579	CCCGGTCAATTTGCTGAGCTGTGATGAGCGCGGCGAGAGCCCTTGACCTTATGAGAGCA	4638
DB	2007	CCCGGTCAATTTGCTGAGCTGTGATGAGCGCGGCGAGAGCCCTTGACCTTATGAGAGCA	2066
QY	4639	CGGGGACCCCTTGGCGGCGACCTTCCACAGAGTTGCCCGTGACCATGCGGAGGCCAC	4698
DB	2067	CGGGGACCCCTTGGCGGCGACCTTCCACAGAGTTGCCCGTGACCATGCGGAGGCCAC	2126
QY	4699	GCGGCGCTGAGAGGCGAGCGCTTTCGTCCAGAGGATCCCGAGACCGGAAAGCTGAC	4758
DB	2127	GCGGCGCTGAGAGGCGAGCGCTTTCGTCCAGAGGATCCCGAGACCGGAAAGCTGAC	2186
QY	4759	GTGAGAGCTCGTGAGATGCGCAAGTCCCGGACAGACACCGTGCCCGAGACCAACCA	4818
DB	2187	GTGAGAGCTCGTGAGATGCGCAAGTCCCGGACAGACACCGTGCCCGAGACCAACCA	2246
QY	4819	CCCATCTTCCCTTATGAGACCTTCTTGGGGCGTGAAGTGCGTGACCTTATGCGAG	4878
DB	2247	CCCATCTTCCCTTATGAGACCTTCTTGGGGCGTGAAGTGCGTGACCTTATGCGAG	2306
QY	4879	CCAGATCCCGCGGCGCTTCCAGCCGCACTCCATCAACCGCGGCGCATCCCTTGAAGGCGAGC	4938
DB	2307	CCAGATCCCGCGGCGCTTCCAGCCGCACTCCATCAACCGCGGCGCATCCCTTGAAGGCGAGC	2366
QY	4939	CGCTGCTACTACCTGCGCGGACACTTGGCGCCCAACCCCACTACCGGCACTTGTATCC	4998
DB	2367	CGCTGCTACTACCTGCGCGGACACTTGGCGCCCAACCCCACTACCGGCACTTGTATCC	2426
QY	4999	ACCTTACTCATCCGCGGCTACCCGACACGCGGCGCTGAGAAACCGGACAGCATCAT	5058

```
|||||
Db 2427 ACCCTACCTATCGCGGGTACCCCGACGCGGCGCTGGAGAACGGGACATCAT 2486
QY 5059 CAATGACTATACATCACTCGGAGAGATGCAACAACGGCCACCGCCATGGCCACG 5118
Db 2487 CAATGACTATACATCACTCGGAGAGATGCAACAACGGCCACCGCCATGGCCACG 2546
QY 5119 AGCTGATATGCTGAGGGGCTCTCGCCCGAGTCTCTGCTGGACTCACTACGTGC 5178
Db 2547 AGCTGATATGCTGAGGGGCTCTCGCCCGAGTCTCTGCTGGACTCACTACGTGC 2606
QY 5179 GGGTCCCCGAGGATCATGCACTGTCTCAAGTGCAACACTGTCTGTGTCTGTC 5238
Db 2607 GGGTCCCCGAGGATCATGCACTGTCTCAAGTGCAACACTGTGTGTCTGTC 2666
QY 5239 GAGACAGAGGACCCCGAGGACCCGCAATGACCCGCTTGTCTTACTCTCCCAACCGGCCCA 5298
Db 2667 GAGACAGAGGACCCCGAGGACCCGCAATGACCCGCTTGTCTTACTCTCCCAACCGGCCCA 2726
QY 5299 GCCCTTCAGAGCGGCAAGAGCTCCCACTCTCCCAAGAGTCAACAACACTTGAC 5358
Db 2727 GCCCTTCAGAGCGGCAAGAGCTCCCACTCTCCCAAGAGTCAACAACACTTGAC 2786
QY 5359 AAAACCAACCAACGCTCTGTCTCGAGCGGAGCGAGACCGGATGAGAGCGGACG 5418
Db 2787 AAAACCAACCAACGCTCTGTCTCGAGCGGAGCGAGACCGGATGAGAGCGGACG 2846
QY 5419 GGAATCGGAGCGGAAAAAGTCTCTCACTGTCAACAACGAGTGAAGACCACTCAT 5478
Db 2847 GGAATCGGAGCGGAAAAAGTCTCTCACTGTCAACAACGAGTGAAGACCACTCAT 2906
QY 5479 CTGGAAGCTGTATCAGAGAGAGAGCGGACGAGCGGAGCGGAGCGGAGGAGG 5538
Db 2907 CTGGAAGCTGTATCAGAGAGAGAGCGGACGAGCGGAGCGGAGGAGGAGG 2966
QY 5539 CAGCAGACGCGCGCCGCTCTCCACTTCCATGCCCACAGCACTTCCCTCTG 5598
Db 2967 CAGCAGACGCGCGCCGCTCTCCACTTCCATGCCCACAGCACTTCCCTCTG 3026
QY 5599 GACCCAGATGCTCTCTCAGAGAGACCCAGTGTCTTCAACAACAGGATGAGAGGTAT 5658
Db 3027 GACCCAGATGCTCTCTCAGAGAGACCCAGTGTCTTCAACAACAGGATGAGAGGTAT 3086
QY 5659 CATCACGCTGTGAGAGCGGACGAGTCTGAGGTCAACTCTCACTCTCACTC 5718
Db 3087 CATCACGCTGTGAGAGCGGACGAGTCTGAGGTCAACTCTCACTCTCACTC 3146
QY 5719 CGTTGCGCAGCTGCAATTCACCTGCACTGCACTGCACTGCACTGCACTGCA 5778
Db 3147 CGTTGCGCAGCTGCAATTCACCTGCACTGCACTGCACTGCACTGCACTGCA 3206
QY 5779 TGGGGTCTACCTTACCTTCAATGAGCGCTGTGCTGCTGCTGCTGCTGCTGCT 5838
Db 3207 TGGGGTCTACCTTACCTTCAATGAGCGCTGTGCTGCTGCTGCTGCTGCTGCT 3266
QY 5839 CGGGCAGAGCGGCGCGGAGAGACCGGCAATGCTTCTGCGCAAGGCCCGCAGCGG 5898
Db 3267 CGGGCAGAGCGGCGCGGAGAGACCGGCAATGCTTCTGCGCAAGGCCCGCAGCGG 3326
QY 5899 CTCGCGGCTGAGAGCGGCTCTCTCCCGCAAGAGGCTCGAGAGCGGCGGCTGAGTGC 5958
Db 3327 CTCGCGGCTGAGAGCGGCTCTCTCCCGCAAGAGGCTCGAGAGCGGCGGCTGAGTGC 3386
QY 5959 TCTGTGCTCTGAGCGGACCACTGCGCGGACCCCTGCGAAGAACTCTGCACTCA 6018
Db 3387 TCTGTGCTCTGAGCGGACCACTGCGCGGACCCCTGCGAAGAACTCTGCACTCA 3446
QY 6019 CGCAGCGCGGAGCGCGCGGCGCACTGCTGCGGCTGCGAGCCGCGCAAGGAAAAAG 6078
Db 3447 CGCAGCGCGGAGCGCGCGGCGCACTGCTGCGGCTGCGAGCCGCGCAAGGAAAAAG 3506
QY 6079 TCAAGTAAACCTTTTCATCCAGAACTGAACTCCGTTCTCTGGGTTACACGCGAG 6138
|||||
Db 3507 TCAAGTAAACCTTTTCATCCAGAACTGAACTCCGTTCTCTGGGTTACACGCGAG 3566
QY 6139 CAGCTACGCGCGGAGAGGGGTGAGCGCGTCAAGCGCTGTGAGCTCAACCACTGACCA 6198
Db 3567 CAGCTACGCGCGGAGAGGGGTGAGCGCGTCAAGCGCTGTGAGCTCAACCACTGACCA 3626
QY 6199 CGAAGAGGGCTCCCAAGCACTGGAAGAGCTGACAGAGCCACTGAGAGGGAGCT 6258
Db 3627 CGAAGAGGGCTCCCAAGCACTGGAAGAGAGCTGACAGAGCCACTGAGAGGGAGCT 3686
QY 6259 GCGGCCAGAGCAGAGCGCGCTGGAAGCTTGGCGGGGAGGCGGCCACTCTCCACACT 6318
Db 3687 GCGGCCAGAGCAGAGCGCGCTGGAAGCTTGGCGGGGAGGCGGCCACTCTCCACACT 3746
QY 6319 GCGGCGCTGCTGAGAGCGAGCGCTGTCAGCGCGGCTCTCAAGCCGCGCGAGGGT 6378
Db 3747 GCGGCGCTGCTGAGAGCGAGCGCTGTCAGCGCGGCTCTCAAGCGCGCGAGGGT 3806
QY 6379 CAAAGTCAACAGCGGGTGTCAACCTTGCCAGCAATCATGAGGTATCAACAGGA 6438
Db 3807 CAAAGTCAACAGCGGGTGTCAACCTTGCCAGCAATCATGAGGTATCAACAGGA 3866
QY 6439 CTACACCGGCAACACCAAGCACTGAGAGCTCAAGCGACCCCTGCGCCCTCTTACT 6498
Db 3867 CTACACCGGCAACACCAAGCACTGAGAGCTCAAGCGACCCCTGCGCCCTCTTACT 3926
QY 6499 CCTGGGGGAGGCTGCGCGCTGCTGAGACCTTCCGCGCGCAACGAGTACCTTACCTCC 6558
Db 3927 CCTGGGGGAGGCTGCGCGCTGCTGAGACCTTCCGCGCGCAACGAGTACCTTACCTCC 3986
QY 6559 GCCCGGAGCATGTGTGCCCGGCGCGGCTCTCCCAAGGAGGAGGAGGAGAGT 6618
Db 3987 GCCCGGAGCATGTGTGCCCGGCGCGGCTCTCCCAAGGAGGAGGAGGAGAGT 4046
QY 6619 TCCAGAGCCAAACAGACGTGCTTGGGTGTGTGAGAGCGTATTGAACCTGTGTC 6678
Db 4047 TCCAGAGCCAAACAGACGTGCTTGGGTGTGTGAGAGCGTATTGAACCTGTGTC 4106
QY 6679 CCCAGCGGAGGGGATAGACGAGGACAGGGCACTCCGGAAGTGTGTATCCCGTGTGTA 6738
Db 4107 CCCAGCGGAGGGGATAGACGAGGACAGGGCACTCCGGAAGTGTGTATCCCGTGTGTA 4166
QY 6739 CCGGATGAGGAGAAACAGACGAGCGCAGAGATGGGCTCCAGTCTTCAGGCAACACAG 6798
Db 4167 CCGGATGAGGAGAAACAGACGAGCGCAGAGATGGGCTCCAGTCTTCAGGCAACACAG 4226
QY 6799 CAGCGCGCACTCTTTCAGAGCTGACCGAGAGCACTCCGCAATGGTCAAGTCA 6858
Db 4227 CAGCGCGCACTCTTTCAGAGCTGACCGAGAGCACTCCGCAATGGTCAAGTCA 4286
QY 6859 GAGGCAAGAGATTCAGAGAGCTGAACCCCAACCCGGAATGAGCTGAAATACATAT 6918
Db 4287 GAGGCAAGAGATTCAGAGAGCTGAACCCCAACCCGGAATGAGCTGAAATACATAT 4346
QY 6919 CAGCGAGCTGAGAGCGAGATCTTCAATATGCGCGCATCAACGGAACAGGCTTATGAC 6978
Db 4347 CAGCGAGCTGAGAGCGAGATCTTCAATATGCGCGCATCAACGGAACAGGCTTATGAC 4406
QY 6979 CTATGAGAGCAGGCGGTGAGGAACATGCAACCAATGAGGCTGAGAGCCATAT 7038
Db 4407 CTATGAGAGCAGGCGGTGAGGAACATGCAACCAATGAGGCTGAGAGCCATAT 4466
QY 7039 TAGAAGGCACTCATAGGTAAATATACAGATGGGAAGGTCCCGCGCTGACGCGCAA 7098
Db 4467 TAGAAGGCACTCATAGGTAAATATACAGATGGGAAGGTCCCGCGCTGACGCGCAA 4526
QY 7099 TGCTTTAAACCTCTGATATGCAAGTCCAGGCTGCGCTGCTATGCTTAAACGCTGC 7158
Db 4527 TGCTTTAAACCTCTGATATGCAAGTCCAGGCTGCGCTGCTATGCTTAAACGCTGC 4586
QY 7159 TGAAGAGCGAGTGAACCACTCACTCTGCGAGGTGCGGCGGAGAGGCAAGGTCTC 7218
Db 4587 TGAAGAGCGAGTGAACCACTCACTCTGCGAGGTGCGGCGGAGAGGCAAGGTCTC 4646
```

QY 7219 TGGCAGACCCAGCAGCCGAAAGCCAAAGTCCCCGGCCCCGGGCTGGCATCTGGGAGCCG 7278  
DB 4647 TGGGAGACCCAGCAGCCGAAAGCCAAAGTCCCCGGCCCCGGGCTGGCATCTGGGAGCCG 4706  
QY 7279 GGCACCTCTGTCTCTCTCATGTGCACTCGGAGGAGAGTGTCAACCGCGGACCGCTCTAC 7338  
DB 4707 GGCACCTCTGTCTCTCTCATGTGCACTCGGAGGAGAGTGTCAACCGCGGACCGCTCTAC 4766  
QY 7339 CAACCGCGGTGGAGAGAGCCCGCTCGGAGGTTCCAGCCATTCCCTTACAACC 7398  
DB 4767 CAACCGCGGTGGAGAGAGCCCGCTCGGAGGTTCCAGCCATTCCCTTACAACC 4826  
QY 7399 CTTGATCATGTGCGGTGTGAGCGGAGTGTCAATGCTTCCCAACCCGACCGGCTCCCGC 7458  
DB 4827 CTTGATCATGTGCGGTGTGAGCGGAGTGTCAATGCTTCCCAACCCGACCGGCTCCCGC 4886  
QY 7459 GGGCAGCGGGCCCTCTGCTGGCCCCCAACCGCTGGAGCGAGAGCCCAAGCCACTGCT 7518  
DB 4887 GGGCAGCGGGCCCTCTGCTGGCCCCCAACCGCTGGAGCGAGAGCCCAAGCCACTGCT 4946  
QY 7519 CTGCTCGAGTACGAGACACTCTCGAGCAGGAGTGACTCAAGAGCGGGGGGGGGC 7578  
DB 4947 CTGCTCGAGTACGAGACACTCTCGAGCAGGAGTGACTCAAGAGCGGGGGGGGGC 5006  
QY 7579 GGGCGGTGERSPRUUGTCAAGTCCAGCGAGCAGCAGAGAAAGCGCTGAGAGCGG 7638  
DB 5007 GGGCGG-----TGTCAAGTCCAGCGAGCAGCAGAGAAAGCGCTGAGAGCGG 5056  
QY 7639 GGGCGGTGCGACTCCCCCAACAGAGAGAGCCCTGAGTCCGCTGCTCCATCCA 7698  
DB 5057 GGGCGGTGCGACTCCCCCAACAGAGAGAGCCCTGAGTCCGCTGCTCCATCCA 5116  
QY 7699 TCTGTCCGTCCAGAGCGGCGCATCTTGCTGTAAAGCTTAAAGCTCCGCCCC 7758  
DB 5117 TCTGTCCGTCCAGAGCGGCGCATCTTGCTGTAAAGCTTAAAGCTCCGCCCC 5176  
QY 7759 GGGCTGCGCTGTGAGACTTACTCAGGGAGTGTTTACTGTGTGTCGAGAGAGAGG 7818  
DB 5177 GGGCTGCGCTGTGAGACTTACTCAGGGAGTGTTTACTGTGTGTCGAGAGAGAGG 5236  
QY 7819 GAAAGGGCGCGGAGAGGCGAGCGAGCGGTGTGCGACACACAGCGCGCCAGGGG 7878  
DB 5237 GAAAGGGCGCGGAGAGGCGAGCGAGCGGTGTGCGACACACAGCGCGCCAGGGG 5296  
QY 7879 GCCAGAGACCCAAAGAGAGTACACAGCAGCTCCAGCGCAGTCCCGCGAATGAT 7938  
DB 5297 GCCAGAGACCCAAAGAGAGTACACAGCAGCTCCAGCGCAGTCCCGCGAATGAT 5356  
QY 7939 TTGGAACCAAGTCTAACTGAGCTGACAGCCCGCGGCTCCCTCGCTCCATCCC 7998  
DB 5357 TTGGAACCAAGTCTAACTGAGCTGACAGCCCGCGGCTCCCTCGCTCCATCCC 5416  
QY 7999 GCTTGCGCTGTGAGACAGTGAAGCGAGGCTGTTCAGCGCCAGTGGCTGTCCG 8058  
DB 5417 GCTTGCGCTGTGAGACAGTGAAGCGAGGCTGTTCAGCGCCAGTGGCTGTCCG 5476  
QY 8059 TCCCAACAGACTGCCCCAGCCAGAGATTTGCGAAGAACAGTCAAGGCGAGTGGGCG 8118  
DB 5477 TCCCAACAGACTGCCCCAGCCAGAGATTTGCGAAGAACAGTCAAGGCGAGTGGGCG 5536  
QY 8119 ACAAAGGGCCAGTGTGCGCTGTGGGGGAGACAGATGCTCCGAGGACTGATTTTTTT 8178  
DB 5537 ACAAAGGGCCAGTGTGCGCTGTGGGGGAGACAGATGCTCCGAGGACTGATTTTTTT 5596  
QY 8179 CACACATGTTTCCCGCAGCGGTGGAGAGAGAAAGCGAGATTAATGATGTGTTGTTAC 8238  
DB 5597 CACACATGTTTCCCGCAGCGGTGGAGAGAGAAAGCGAGATTAATGATGTGTTGTTAC 5656  
QY 8239 AGGATATATTTTGAATCCTTCAATGAATTAATTCAGATGTTTTCGCAAGAGAGACT 8298  
DB 5657 AGGATATATTTTGAATCCTTCAATGAATTAATTCAGATGTTTTCGCAAGAGAGACT 5716

QY 8239 ACCGAGTATTACTGCTGTGCTGCTTTTGAATCTCTGCTTACCGTTCAAGAGCGGTGTGAG 8358  
DB 5717 ACCGAGTATTACTGCTGTGCTGCTTTTGAATCTCTGCTTACCGTTCAAGAGCGGTGTGAG 5776  
QY 8359 GCCGACATTCGTGAGCCCATCATCTCGAGAGCAAGGGGGGGGAGCTGCTCGTACGC 8418  
DB 5777 GCCGACATTCGTGAGCCCATCATCTCGAGAGCAAGGGGGGGGAGCTGCTCGTACGC 5836  
QY 8419 CCCGCTGTGCTCTCCCTCCCTCCCTCCCTTCCCTGGGAGAAATGAATGATGCGTATTTCTGTG 8478  
DB 5837 CCCGCTGTGCTCTCCCTCCCTCCCTCCCTTCCCTGGGAGAAATGAATGATGCGTATTTCTGTG 5896  
QY 8479 GCCGCAATTTGGCAGAGGTGTGTATTTCTGTCATTTAACAACAGCTGTTCTAATTAATAA 8538  
DB 5897 GCCGCAATTTGGCAGAGGTGTGTATTTCTGTCATTTAACAACAGCTGTTCTAATTAATAA 5956  
QY 8539 GCGAATTAATCTCCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 8571  
DB 5957 GCGAATTAATCTCCAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5998

RESULT 6  
AF125671  
LOCUS  
DEFINITION  
Mus musculus silencing mediator of retinoic acid and thyroid hormone receptor extended isoform (Smrte) mRNA, complete cds.  
ACCESSION  
AF125671  
VERSION  
AF125671.1 GI:459295  
KEYWORDS  
Mus musculus (house mouse)  
SOURCE  
Mus musculus  
ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.  
1 (bases 1 to 8544)  
Park, E.J., Schreien, D.J., Yang, M., Li, H., Li, L. and Chen, J.D.  
Smrte, a silencing mediator for retinoid and thyroid hormone receptor-extended isoform that is more related to the nuclear receptor corepressor  
Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3519-3524 (1999)  
1097068  
JOURNAL  
PUBMED  
REFERENCE  
2 (bases 1 to 8544)  
Chen, J.D.  
Direct Submission  
Submitted (03-FEB-1999) Pharmacology and Molecular Toxicology, University of Massachusetts Medical School, 55 Lake Avenue North, Worcester, MA 01655, USA  
FEATURES  
source  
1..8544  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
1..8544  
/gene="Smrte"  
160..7548  
/gene="Smrte"  
/function="transcriptional corepressor"  
/note="SMRTE"  
/codon\_start=1  
/product="silencing mediator of retinoic acid and thyroid hormone receptor extended isoform"

TANSQGRKGRITRSMANBANHEBTATPOOSSELAEMENNESSRWTEBEMETAKGDL  
EHRNMSALARMVGSKTYSQCKNIFYNYKQRLDEILQOHLKMKERKAKKKT  
PAAASEBETAPPAABDEMBASGANSANEELEABEASASQSEVREPGSCAPAVN  
NSDTPESVSPRESEATQDQPKPTQTEALPAATQPPPREBAPAAAPABSPDASG  
PPSPESPSPAPAPATVDKQBOAPAPAPOTEDAKQYSBAEITDVQPREPEABEE  
PPESVSDKHEETEEBPEDEKAKTEALIVTSEAPLKEBEGSAVAATVKSSSGATDS  
DSATCSABEVEBEGDGKRLSPRPSTLTPAGDPRASTSPKQDLKQKORAAI  
PIVTVHEBREDVTVPKPVPPVPTQHLQEGDVSQOOSGSPPGSKRSVPYPAEK  
BAEKPAFPAPPTBEGOSYRLSPHAGRLBSPREVIKTSTRADPLPSTYTPHNPPL  
GLHDSAPVLPAPPPIISSAKHPCVLEBROGLATIQGMSVCLRVHSEBAKMG  
PLTMELPLAVDPKCLSTALAPPEVASPPASQTPGCRPOLORLHPRTPADVLKGT  
ISRTVGBDSPSRLDRABEDTLPKGNVLYEGKKGAVLSYEGMSVSCSKDEGSSSGP  
PHETAPKRTYDMMGEVGTYSASIEGMLKRAIPEOSPHLEKHNIIRGSTTQOIP  
RSVYAEDEYLREREAALKREBTPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP  
ATYKAGRSIHIEPREELRRTPELPLAPPLKSGSTTQOPLKYDSAGSTQKKNV  
REIISGRPRPRALHPDLMADRALERACYEBSLSKRSSTSGAGSITTRGAPVVP  
ELKRPQSPUTYEDHAPFTSHLPKSPVATREPTVRLDEGLSLSKASQDRKLSTP  
BIKASPHSTVEHNPPIISPYEHLRGVGLDYRGLPLAEDPISIRGJPLEMAA  
AAVYLRHLAPSTYPLHYPLRYGPTALLENQTIINDYITISQOMHNAASMA  
QADMLRGLSPRESSLALNYAGPRGIILDSQVPHLPVLPVPTGPATADBLAYLP  
TAPRESSHSSPLSPGGPTHLAKPTATSSSRERERENDKSLITTTTVEHAP  
WMPRTSSSGAGSSRPASHTHOHPISRTQDALOQPRSVLANTMKGVTSVEGCT  
PVLNISTSSPVRPAPATPPATNCPUGGLBGLVPLTMEVPLPRETSRVAPBEPK  
VYGAFLTKPAPAREPASPSKSSBRSJLAPSSHTALARTPAKSLAPHASBPBG  
PTASDLHREKTQSKFPIQELERSLGTHSAGYSPDVEPISPVSSPLTHDKLS  
KLEBELKSHLEBELKHOQPMKLSABAHLPHLPBPESOPSSPLLOTAGIKH  
QEVVTLAHSIETVITQDTRHHPQOLSGPLPALVSPASCVNLDLRPSQLYLP  
PHGTPARSGPSHSEGGKSPSPBKTSTVLSSEDAIIPVPPREMTBPGHARSAYTL  
YNDGEGEPRMGSKSPGNTSQPPAPFSKLTESNAVAKSKOEINKLANTHNNBEPY  
LNGOQGETEIPNMPATITAGIMTCRQAVOVBHASTNMGLRAIIRKALMGYTDWNEBP  
LGANAFNPLNASASLPAAMPITTAGDRBDHALTSPGGGAKVAPSPRSRKAQVAP  
GLASGDRPSPVSVHSEDPQNRRTPLTNVWEDRPSASTPPVPLMRLQAGVMA  
SPPPEPLAAGSPHAGPHHAMBEPKPLCSQYETISDBE"

ORIGIN

Query Match 58.8%; Score 5037.2; DB 9; Length 8544;  
Best Local Similarity 78.3%; Pred. No. 0;  
Matches 6770; Conservative 0; Mismatches 1493; Indels 387; Gaps 45;

QY	1	CATGTCGGGCTCCACACACTTGTGGCACAAGAGTGAAGGCGCACTGAGCCCGGCTAACCC	60
DB	159	CATGTCAGATCCACACACCTGTGGCACAAGATGGCGGGCTGTGAAGCCCGGCTAACCC	218
QY	61	GGCCCAAGCCTTCTTACCCAGTGCAGATCGCCGGAGCGCACAGGAGCGTGGGCTGCT	120
DB	219	ACCCCAAGCATCTCTTACCCGCTGCAATAGCCCGGCTCCACAGGAGCGTGGGCTGCT	278
QY	121	GGAGTACAGACCACTCCCGGACTATGCTCCCACTGTGCGCGGCTCCATCATCCA	180
DB	279	TGAGTACCAACACACCCCGTGACTACACTCAGACTGTGACCGGGTTCCATCATCCA	338
QY	181	GGCCCAAGCGGCGAGGCGCTCCCTGCTGTGTGAGTTCAGACC GGGAATGAACGTTCCA	240
DB	339	GGCACAAGAGAGGCGGCGCTCACTGCTGTCAAGATTCCAGCTGGAGTGAACGTTCTCA	398
QY	241	GGAGCTCCACCTGCGGCGAGAGTCCCATATPCCGCGGAGCTGGGGAAGTCAAGAT	300
DB	399	GGAGCTCCACCTGCGGCTGAGTCCCGCAGCTTCGCTGAGACTGGGCAAGCCGACAT	458
QY	301	GGAGTTCAATTGAAGCAAGCGCCTCGGCTAGAGCTGTGCTGACCCCTGTGCGACC	360
DB	459	AGAAATTCACGAGAGAGAGCGCCCGGCTGGAGCTACATACCGGATACCTGTGTGCGCC	518
QY	361	GTTCACCCCTGTGCGCAGCGGCGAGCTTGCGGAGTTGAAGCTTCAACAGGACCTGAG	420
DB	519	ATTCACCCCTGTGCGCAGCTGGGCGAGCGAGTGGTGTGAAGACTTAAACAGGACCGTAG	578
QY	421	CGTGAACGGGCAAGCTGGAACCGGTGTCTCCCGCAGCCCGGCAACTGAACCTGAGCT	480
DB	579	CTTGGGAGGCAAGCTTGAAGCTGTGTCACTCCCAAGTCCCGCAGGCTGACCTTGAAGCT	638
QY	481	GGAGCTGTGCGGCGCAAGCTGTCCAAAGAGGAGCTGATCCAGAAACA---TGAGCCGCGT	537

DB	639	AGAGCTGGCGCCATCTGCACTGTCCAAAGAGAGCTGATCCAGAAACAATTTGACCGCGT	698
QY	538	GGACCGAGAGATCAACATGGTGTAGAGAGAGATCTCTAAGCTGAAGAAAGAGCAGACA	597
DB	699	GGACCTTGAGATACCATGGTGTAGAGAGAGATCTCCAGCTGAAGAAAGAGCAGACA	758
QY	598	GCTGAGAGAGAGGCTGCCAAGCGCGCCGAGCCTGAAGAACCCGTGTCAACCGCGCCAT	657
DB	759	GTTGAGAGAGAGGCGCGCCAGCGCGCCGAGACCCGAGAGCCTGTGTCCGACACCCAT	818
QY	658	CGAGTGAAGAGCAGCGAGCTGTGTGAGATCATCTTCAGACAGAGAACCGGAAGAGCTGA	717
DB	819	AGATTAAGAGCAGCGAGGCTGTGTCCAGATCATCTTCAGATGAAGAACCGGAAGAGCGCA	878
QY	718	AGCTGACATCGGATTTCTGAAGAGCGCTGGGGGCCCGAGGTGAGACTCCGCTGTACAAACA	777
DB	879	AGCGGACACCGGATCTTGAGAGGCTGGGGGCCCGAGGTGAGACTCCCTGTGTACAAACA	938
QY	778	GGCCTCCGACACCGGCAAGTATCATGAGAACATCAAAATTAACCAAGCGGTGCGAAGAA	837
DB	939	GGCGTGTACACACGCGCAGTACATGAATAACATCAAAATTAACCAAGCGGTGCGAAGAA	998
QY	838	GCTAATCTTGTACTTCAAGAGAGATACGCTCGGAAAACAATGAGAGCAAGTTCTG	897
DB	999	GCTGATCTTGTACTTCAAGAGAGATACGCTCGGAAAACAAGTTGAGAAACAGCGCTTCTG	1058
QY	898	CGAGCGCTTGTACAGCTCATGAGAGCGCTTGAAAAAAGGTGAGACCGCATGCAAAAAACA	957
DB	1059	CGAGCGCTTGTACAGCTCATGAGAGCGCTTGAAAAAAGGTGAGACCGCATGCAAAAAACA	1118
QY	958	CCGCGCGCGCGGCGCAAGAGAGCAAGGTGCGCGAGTACTAGAAAAGCATTTCCCTGA	1017
DB	1119	TCCGCAAGAGAGGCGCAAGAGAGCAAGGTGAGAGTACTAGAAAACAGTTCCCGGA	1178
QY	1018	GATCCGCAAGAGCGGAGCTGTGAGAGCGCATGCAAGACGGGTGGCCAGCGCGGCGAG	1077
DB	1179	GATCCGCAAGAGCGGAGCTGTGAGAGCGCATGCAAGACGGGTGGCGAGCGGTGGCGAG	1238
QY	1078	TGGGCTGTCAATGTGCGCGCGCGGAGAGCAAGAGGTGTCAGATCATGATGATGGCCCT	1137
DB	1239	TGGGCTGTCAATGTGCGCGCGCGGAGAGCAAGAGGTGTCAGATCATGATGATGATGGCTT	1298
QY	1138	CTCAGAGCAGAGAACTGTGAGAGAGATGCGCCAGCTGTGCGCGGCTCCAGTGT	1197
DB	1299	GTCTGAGCAGAGAACTGTGAGAGAGATGCGCCAGCTGTGCGCGTATCC--GCCATGTT	1356
QY	1198	GTACGAGCTGTACACAGAGCGCATCAAGTTCAATCAACATGAACCGGCTTATGGCCGACCC	1257
DB	1357	GTACGAGCGC-GACACAGCAGAGATCAAGTTCAATCAACATGAATGACTATGGATGACCC	1415
QY	1258	CATGAAGGTGTCAAAAGACCGCCAGGTCAATGAACATGTGAGGTGAGCAAGAAAGAGAC	1317
DB	1416	CATGAAGGTGTCAAAAGACCGTCAGGTTCACAAAGTGAAGAGCAAGAGAGGAGACAC	1475
QY	1318	CTTCCGGAGAGATTCAATGACAGATCCAAAGAACTTTGGCTGATGCAATCTCTGGA	1377
DB	1476	CTTCCGTGAGAGATTATGACAGACCTTAAGAACTTTGGCTGATGCTCATTTCTGGA	1535
QY	1378	GAGGAAGACAGTGTGAGTGTGCTCTATTAATTAAGTGAAGAAATGAAGACTA	1437
DB	1536	GAGGAAGACAGTGTGAGTGTGCTCTATTAATTAAGTGAAGAAATGAAGAAATTA	1595
QY	1438	TAAAGCTGTGTGAGACGGAAGCTTATGTGGCGCGGCGCAAGACCGACGACCAACAAGCA	1497
DB	1596	CAAGACTTGTGTGAGGCGGAGCTATGTGGCGCGGTGCAAGAC-----CAGCA	1643
QY	1498	GCAG	1557
DB	1644	GCAG	1703
QY	1558	GAAAGTGAAGAGAGAGAGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1617
DB	1704	GAAAGTGAAGAGAGAGAGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1763

QY 1618 GAAAGCAAGAAAGACTCTCTCAAGAGAAAGACAGACGACCTCGAGGGAGAGAACAG 1677  
DB 1764 GAAAGCAAGAAAGAACTAGAGAGAGAGAGACAGACGACCTTCTGGAGAGAGACAGAG 1823  
QY 1678 CGAAGAAAGAGCTGTGGCTTCCAAAAGCCCGCAAACTGCGCAACGACGAGGAAAGCGCA 1737  
DB 1824 TGAGAAAGAGGCGGTGGCTTCCAAAAGCCCGCAAACTGCGCAACGACGAGGCGCGCGCA 1883  
QY 1738 AGGCGGCATCACCGCTGCAATGGCTTAATGAGGCGCAACGCGAGGAGGCGCATGCCGCCA 1797  
DB 1884 AGGCGGTATCACGCGCTCATAGCGCAAGAGCGCAATGAGAGACAGCCACCCACA 1943  
QY 1798 GCAAGAGCCGCGAGCTGAGCTCCATGAGAGCTGATGAGAGTTCTCGCTGAGCAGAAAGAG 1857  
DB 1944 GCAAGATTCAAGAGCTGGCTTCCATGAGAGATGAAACGAGAGTTCTCGCTGAGCTGAGAGAG 2003  
QY 1858 AATGGAACAGCGCAAGAAAGGTCTCTGGAACAAGCGCGCAACTGGTCGGCATGCGCG 1917  
DB 2004 GATGAGACAGCAAGAAAGAGCGCTCTGGAAATGAGAGAACTGGTCAGCATTTGCCG 2063  
QY 1918 GATGGTGGGCTCCAGACTGTGTGCGAGTGAAGAACTTCTAATTCACTAGAGAGAG 1977  
DB 2064 CATGGTGGGCTCCAGAGCCGTGTCCAGTGTAAAGCTTCTACTTCACTAGAGAGAG 2123  
QY 1978 GCAGAACTCGATGAGATCTTGCAGAGACACAAGCTGAAGATGAGAGAGAGAGAGAG 2037  
DB 2124 GAGAAACCTCGAGAAATCTTTCAGAGACACAAGCTAAAGATGAGAGAGAGAGAGAG 2183  
QY 2038 GCGAGAAAGAAAGAAAGCCGCGCGCGCGCGCGCGAGAGAGAGCTGATTCGCCCGCT 2097  
DB 2184 TCGAGAGAAAGAAAGAAAGCCGAGCTGCGCGCGAGCGAGAGAGAGCTTCCCACTGCG 2243  
QY 2098 GGTGAGAGATGAGAGATGAGAGCGGTGCGCGCGTGAAGCGGAAATGAGAGAGAGATGGTGA 2157  
DB 2244 GCTGAGAGAGAGATGAGAGATGAGAGATGAGAGCGAGCGAGAGTGCATGAGAGAGAGCTGGCG 2303  
QY 2158 GAGAGCTGAAGCTTACATGCTCTGAGAGATGAGAGTGGCCAGA---GGGAAATGAGAGTGG 2214  
DB 2304 GAGAGCAAGAGCTCAAGGCGCTCTGAGAGATGAGAGTTCCTCAAGATTGGGAGATGAGAGTGG 2363  
QY 2215 CCGAGCACTGTCAACACAGCTCAGACACCGAGAGATCCCTCTCTCAACCTGAGAGC 2274  
DB 2364 CCGAGCTGTGTCAACACAGCTGTGATCTGAGAGTGTCCATCCCGCTTCCAGAAAGC 2423  
QY 2275 GCGCAAGAGCAAGAGGCAAGATGGGCGCAAGGCCCGCAGCACTGAGGAGCGCGAGCGGC 2334  
DB 2424 CAGGAAGGACAC-----TGGGCTTAACCCCACTGGCACTGAAGCATTTGCCGCTGC 2474  
QY 2335 ACCCCAGGCCCCACCCACCCACACGAGAGCATCCCGGAGCCCCCATTTGAGAGCCACCC 2394  
DB 2475 CACCGAGCACTGTTCCTCTCTCAAGAAAGACCGGAGAGAGCCCTGCTGAGAGCCCTCCCC 2534  
QY 2395 GGCCTGTGAAGCAACCGAGAGCCCTTAAGCGCCCCCACCAGACACCCCATGCGCTCTGACCC 2454  
DB 2535 AGTCCCTGTATGCA---GTGGCCCAACATCCCAAGAGCTTCCCATCACTGCGCGCAC 2591  
QY 2455 TCCTCTGTGTGCTCCCAAGTGA 2514  
DB 2592 CCGGCTACTGTGAGACAG 2651  
QY 2515 GAGAGGAG 2556  
DB 2652 AGATGCGCAAG 2711  
QY 2557 AGTGAACACAG 2616  
DB 2712 CGAGGCTCTGTGAG 2771  
QY 2617 GGGAGCGAGC---CAAG 2673  
DB 2772 AGAGCTGAAGACAAAG 2831

QY 2674 CAAGCAGAGAAAG 2733  
DB 2832 TAAAG---GTGAGAGAGAGCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2888  
QY 2734 CCCCCAG 2793  
DB 2889 CACCGAG 2948  
QY 2794 CCGGCAAG 2853  
DB 2949 AGGTGAACAG 3008  
QY 2854 CCGGCGCAATGCTTCAACCCAG 2913  
DB 3009 CCGGCGCAATGCTTCAACCCAG 3068  
QY 2914 TGCATTCCTCCCATTCAGATTCACAAAGTCCATTAAGAGAGAGAGAGAGAGAGAGAGAG 2973  
DB 3069 GCGCATCCCCCTATC---GTGACCAAGAGTCCATTAAGAGAGAGAGAGAGAGAGAGAGAG 3125  
QY 2974 CACCAAGCAGAGCTCCCCAG 3033  
DB 3126 CCGCAAGCAGAGTCCCTCTGTGCTTCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3185  
QY 3034 CCTTCAGAGAGCTGAG 3093  
DB 3186 GTCTCAGAGAGTGGAG 3245  
QY 3094 CAAGAGAGC-----CTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3130  
DB 3246 GAAAG 3305  
QY 3131 GGGAG 3190  
DB 3306 ACTGAG 3360  
QY 3191 AAGGCTCCCGAGATGCCCCGAG 3250  
DB 3361 AAGACTTCACACAGAGC-----TGAACCTCTCTTCTCTCAACAGAGAGAGAGAGAGAG 3414  
QY 3251 CCGGCGCGAG 3310  
DB 3415 CCGGCTGTGGGCTTCAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3471  
QY 3311 AACCGGCTCCCGCTCATCTCTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3370  
DB 3472 AACCCCAAG 3531  
QY 3371 GCGATCTCCCAAG 3430  
DB 3532 GCGATCTCCAG 3588  
QY 3431 CCGGAG 3490  
DB 3589 CCGAG 3642  
QY 3491 TTCAAG 3550  
DB 3643 ----- 3642  
QY 3551 CTGGGAG 3610  
DB 3643 -----GGGACAG 3662  
QY 3611 CCGGAG 3670  
DB 3663 ACGAGTGAAGAGATACAG 3715  
QY 3671 ACATACCGAG 3730  
DB 3716 AGCTACAG 3774  
QY 3731 ACAGAGATCATGAG 3790



```
Db      3775 AGCAGATCGTCGTGAGGACAGACCCAAATCGCTTGAACGGGACAGAGAGAACCTTG 3834
Qy      3791 CCCAAGGGCCACGTCATCTTACGAAGGCAAGAGGCCACGCTTTGCTCTTATGAGGGTGC 3850
Db      3835 CCCAAGGGCCATCTCATCTTATGAGGGCAAGAGGCCACGCTCTTCTCTTATGAGGGTGT 3894
Qy      3851 ATGTCTGTGACCGCAGTGTCTCAAGAGAGCGGCGAAGAGAGCTCAGGACCCCCCATGAG 3910
Db      3895 ATGTCCGTGTCAAGTGTCTCTTAAAGAGAGTGAAGAGAGAGCTCTGGGCCACCCCATGAG 3954
Qy      3911 ACGGCGCGCCCAAGCGCACCTTATGACATGATGAGAGGGCCGCGTGGGCAAGGCCATCTGC 3970
Db      3955 ACTGCGCGCCCTAAAGCAGCCTATGACATGATGAGAGGGCCGCTGTGAGGCAAGCTGTACC 4014
Qy      3971 TCAGCCAGATCGAAGTCTCATAGGGCCGTGGCCATCCCGCGGAGGCAACAGCCCCAC 4030
Db      4015 TCAGCCAGATTAAGAGAGCTCATAGGGCCGCGCATCCCTGAGCAGC-----ACAGCCC 4068
Qy      4031 CACCTCAAGAGAGCAGCAGCATCCGCGGGTCCATCAACAAGGATCCCTCGGTCTTAC 4090
Db      4069 CACCTCAAGAGAGCAGATCACTCCAGAGCTCATCAAGAGCATCCGAGGTCTTAT 4128
Qy      4091 GTGAGAGCACAAGAGACTACCTGCGTGGAGAGCCAAAGCTCTTAAAGCGGAGGCGACG 4150
Db      4129 GTGAGAGCGCAGAGAGACTTACGGCGGGAGGCCAAGCTTTGAAGCGAAGAGGACA 4188
Qy      4151 CCTCGCGCCCAACCGCCCTCAAGGAGCTGAACGAGCCTTAAAGCGGAGGCGACG 4204
Db      4189 CCACTCTCCCAACACCACTCTGGAGCTTGAAGCTTAAAGCGCCCGCCCTTGAAC 4248
Qy      4205 ---CTGAGCGCCCTGAAGCTGAAGCGGCGCCATGAGGGCCGTGGTGGCAAGTGAAGAG 4261
Db      4249 CCTCTGGGTCTCTTGAAGCTGAAGCGGACTTCAAGAGGGTGTGGTGAAGCTTGAAGAG 4308
Qy      4262 GCGGCGCGCTCATCATGATATCCGCGGAGAGCTGCGGACACGCGCGAGCTTGCC 4321
Db      4309 GCGGCGCGCTCATCATGATATCCGAGAGAGAGCTGCGGACACCTGAGCTTACCC 4368
Qy      4322 CTGGCCCCCGCGCGCTCAAGAGAGGCTCATACGAGAGGACCCCGCTCAAGTACGAC 4381
Db      4369 CTGGCAACAAGGCTCTGAAGAGGGTTCATCACCAAGGACCCCACTCAAGTACGAC 4428
Qy      4382 ACCGAGCGCTCAACACTGCGTCAAAAAGCAGAGGTACGCTCCCTCATCGGACAGCC 4441
Db      4429 TCTGGAGCACTCTCACTGACACAAAGAACAGAGTGCGCTCATCATCGGAGAGCC 4488
Qy      4442 GGCAGGACGTTCCACCCGTCACCCGCTGATGTATGACGACGCGCGGACCTTGAA 4501
Db      4489 GGCAGGACCTTTCCCTGCGCTGACCGGCTGACATATGCTGACCCCGGACACTGGAG 4548
Qy      4502 CGTGCCTGTCAAGAGAGAGCTGAAGAGCGGCGCAAGGACCGGACAGACTCGGAGGCG 4561
Db      4549 CGTGCCTGTATGAAGAGAGTGAAGAGCGGCTCAGGAGCAGCAGTGTGACGGGGCG 4608
Qy      4562 TCCTATGCGCGGCGCGCGCGCTCATTTGCGCTGAGCTGAGTAAAGCGCGGACAGCC 4621
Db      4609 TCCTATACAGTGGGGCTCAAGTGTGTGCTGAATGAGGCAAGGCAAGCAAGCCCA 4668
Qy      4622 CTGACCTATAGAGACACAGGAGGACCTTTGCGGACACTCCCAAGAGTTGCGCGTG 4681
Db      4669 CTGACCTTAAGAGACACAGGAGGACCTTTCAACAGTCACTGCAAGTGTCTCCCTGTG 4728
Qy      4682 ACCATCGGAGAGCCACGCGCGCGCTGACAGAGGAGAGCTTTGTCAGAGAGCACTTC 4741
Db      4729 ACCACAGAGGAGGCGCACGCGCTTCAAGAGAGGAGGAGCTCTTATCAAGCAAGGGGTC 4788
Qy      4742 CAGGACGAAAGCTGACGTGACGCGCTGAGAGTGCACAAAGTCCCGGACAGCAAGCGTG 4801
Db      4789 CAGGACGGAAGCTGACATCTTACACCCCGGAGAGTGCACAAAGTCCCAACAGCACTGTG 4848
Qy      4802 CCGAGACACACCAACCCCATCTGCGCTTATGACACCTGCTTGCGGCGCTGAGTGGC 4861
Db      4849 CCCAGACACCAACCTTCACCCCATCTCCCTATAGACCTTGTCCGGGCGTGAAGTGT 4908
Qy      4862 GTGACCTGTATGAGCAAGCCACATCCCGTGGCCCTTGAGCCCACTTCATACCCCGGCG 4921
Db      4909 GTGACCTGTATCCGTGTACATCTTACATGAGCTTTGATCCCACTTCATACCCGAGGG 4968
Qy      4922 ATCCCTGTGA---CGAGCGCGCTGACTTACCTGCGCGGACAGCACTGAGGCCCAACCC 4978
Db      4969 ATCCCTGTGAACAGACAGCGGAGCTTACTACTGCGCGGACACTTGGCCCCAGGCC 5028
Qy      4979 ACCTACCCGACCTGTATCCCACTTACCTTATCGCGGCTACCCGACAGCGCGGCGTG 5038
Db      5029 ACCTACCCACACCTGTATCCCACTTACTTATCGCGGCTACCTTACAGCGCGGCGCTG 5088
Qy      5039 GAGAACCGGACGACATCAATCAATGACTATCATACCTCGAGAGATGACCAACAGCG 5098
Db      5089 GAGAACCGGACGACATCAATCAATGACTATCATACCTCGAGAGATGACCAACAGCT 5148
Qy      5099 GCCACCGCATGAGCCCAAGGAGCTGATATGCTGAGGGGCTCTGCGCCGCGAGTCTCG 5158
Db      5149 GCTCCGCGCATGAGCCCAAGGAGTGTGACATGTGAGGGGTGTGACCGGAGAGTCTCG 5208
Qy      5159 CTGGCACTAATCACTGCTGCGGCTCCCGAGGCAATGACCTGTGCCAAGTGCACAC 5218
Db      5209 CTGGCCCTCAATTAATCCGCTGGCCCAAGAGGCAATATGACCTGTGCCAAGTGCACAC 5268
Qy      5219 CTGGCTGATGTGTGCGCCCGACACAGAGCAACCCGACCGGCAATGACCGGCTTGGC 5278
Db      5269 CTGGCTGTGTGTGTGACCAACAGCGGACGCGCAACCGGCAATGACCGGCTTGGC 5328
Qy      5279 TACCTCCCAACGCGCGCCGACGCTTCAAGAGCGGCAAGAGAGTCTCCCACTTCCCA 5338
Db      5329 TACCTCCCACTGCGGCCCAACCTTCAAGAGCGGCAAGAGTGTACCGGTGTCCCA 5388
Qy      5339 GAGAGTCAACACACTTGAACAAACCAACCAACGATCTCTGTCCAGCGGAGCGAGAC 5398
Db      5389 GAGAGCGCCCACTCACTAGCTAAACCACTGCACTTTCATCGAGCGGAGACG---- 5444
Qy      5399 CGGAGTCAAGAGGAGGACCGGAGTGGGAGCGGAGAAAGTCTCTTCACTTCAACACG 5458
Db      5445 -----GGAAGTGAAGCGGAGAACGAAAGTCACTTCACTTCACTTCACTTCACTT 5490
Qy      5459 ACGGTGAAGACGACACCAATCTGAGAGCCTGTGAAGAGAGAGAGAGAGAGAGAGAG 5518
Db      5491 ACGGTGAAGACGACACCAATCTGAGAGCCTGTGAAGAGAGAGAGAGAGAGAGAGAG 5550
Qy      5519 AGCAGCGCGGGGTGGGGGACAGACAGCCGCGCTTCCACTTCCATGCGCACAG 5578
Db      5551 AGCAGCGG-----CCCGGCTTCCCAACCCACAG 5580
Qy      5579 CACTGCGCCATCTCCCTGAGACCCAGATGCCCTTCAAGACAGACCCAGTGTCTTAC 5638
Db      5581 CACTGCGCCATCTCCCGCGGAGCCGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 5640
Qy      5639 AACACGAGATGAAGGATATCATCAACGCTGTGAGAGCCAGCAAGGCCACGCTTGAAG 5658
Db      5641 AACACGAGATGAAGGAGCTGTCACTTCTGTGAAACCGGACAGCGCTCTGAGG 5700
Qy      5699 TCACTTCACTCTTCACTCACTGCGTTCAGCTGCAATTCCTCACTTCCACCTGAC 5758
Db      5701 TCACTTCACTCTTCTGCTGCTGCGCCAGCTGCAATTCCTCACTTCCACCTGAC 5760
Qy      5759 CCACTGAGGCGGACCTTCAATGAGGCTTACCTTACCTTCAATGAGAGCGGCTTGTGCGCC 5818
Db      5761 CCACTTGTGAGGACCTTGAAGGGGTCTACCTTACCTTATGAGAGCGGCTCTGTAAACC 5820
Qy      5819 AAGAGGCGCCCGGGTGCAGGCGGACAGAGCGGCGCGGAGAGACAGCCGCAATGCTTGC 5878
Db      5821 AAGAGAGCTTCTGCGGTGCGCCGCGGAGCGGCGGCGGAGCGGCTGTGAGAGCTGTGCTTC 5880
Qy      5879 CTGCGCAAGCCCGGCGGCTTCCGGGCTGAGAGCGGCGCTTCTTCCCGACAGCAAGGCTCG 5938
Db      5881 CTCACCAACCCCGGCGG-----GAGAGCGGCTTCTTCAACCGCAAGAGCTCC 5931
```



QY	5935	GAGGCCCGGGCCCTTAATGTCCTTCCTGTCTGTGGCCACGCAACATGCTCCCGACCTTCGCG	5998
Dp	5932	GAGCCGCCATCTCCATGACACCCCCCAGCTCCAGCCACACAGGCAATGCTCCCGACCCACGCA	5991
QY	5999	AAGAACTCGACCTCACACGCGCAGCCCGGACCGCGCGCCACCTGCCTCGGCTCG	6058
Dp	5992	AAGAGCTTGGACCCCACTACATGCAATGCTCCGAACTCCGCCGGGG----CCACCTCGGCTCA	6048
QY	6059	GACCCGCAACCGGAAAGACTCAAGTAACTCTTTTCATCCAGAACTGAACTCCGT	6118
Dp	6049	GATCTGCACCCAGAAAGACTCAAGTAACTCTTTTCATCCAGAACTTGAATCTCGT	6108
QY	6119	TCTCTGGGTTTACAC--GCGACAGCTACAGCCCGGAGGGGGTGGAGCCCTGACGCT	6175
Dp	6109	TCTCTGGGTTTACACAGTGGAGCTGGCTACAGCCCGGATGGGGTGGAGCCCATAGCCCG	6168
QY	6176	GTGAGCTCACCACTGTGACCCACGACAGAGGGGCTCCCAAGCACTTGAAGACTCGAC	6235
Dp	6169	GTGAGCTCCCCCAGCTGTGACCCACGACAGGGGGCTCTCAAACTCTTGAAGACTGAG	6228
QY	6236	AAGAGCCACTTGGAGGGGGAGCTCGCGGCCAAGCAGACCGCCCTGTGAACCTTGGCGG	6295
Dp	6229	AAGAGCCACTTGGAGGGGGAGCTCGCGGCAAGCAGACCGCCCATGAACCTCAGCGCG	6288
QY	6296	GAGGCGCGCCACTTCCACACTGCGGGCGGCTGCTGTGAGCCAGACCTCTGCAGCGG	6355
Dp	6289	GAGGCTGCGCACTTCTCCACATCTGGGGCACTGCGCCAGACCGACCTTATCAGCCCA	6348
QY	6356	CTGCTCCAGACCGCCCGGAGGTCAACAGCGGGTGTGCATCCTTGGCCAGAC	6415
Dp	6349	CTCTCTCCAGACTGCGCCCAAGGACTCAAGAGGTCAACAGAGGGTGTACCTCTGCTAGAC	6408
QY	6416	ATCAGTGAAGTCAATCACAGAGACTACACCGGGACCAACCCACAGAGCTGAGGCAACC	6475
Dp	6409	ATCAGCAAGTCAATTCAGAGACTACACCGGACCAACCCGACAGACTGATGAGCTCC	6468
QY	6476	CTGCCCGCCCCCTTACTCTCTTCCCTTGGGGCAAGCTGCCCCGTCCTTGGACTCCGCGC	6535
Dp	6469	CTTCCCGCCCCCTTACTCTCTTCCCGGAGCAAGCTGCTGTGATGATCTTCCGCGC	6528
QY	6536	CCACCCAGTGAACCTTACCTCCGCCCCCGGACCAATGATGCCCCCGGCTGCTCCCC	6595
Dp	6529	CCACCCAGTGAACCTTACTCTTCCACCCCCCACAATGGCACTCCAGCCCCGGAGATCCCC	6588
QY	6596	CACAGCGAAGGGGCAAGAGGTCTTCCAGAGCCAAACAGAGTGGTCTTGGGTGTGT	6655
Dp	6589	CACAGTGAAGGGGGCAAAAGGTCCCGAGAACCCAGCAAAACATGGGTCTTGGGCAAGT	6648
QY	6656	GAGGACGATATTGAACTGTGTCTCCCAACCGAGGGCATGACGAGGCCAGGGCACTCCCG	6715
Dp	6649	GAGGATGCAATTGAGCTGTGTCTCCCAACAGAGGGCATGATGAGCCAGGACATGCTCG	6708
QY	6716	AGTGTGTGTATCCGCTCTGTGACCGGGATGGGGAAACGACGAGGCCAGAGGATGGGC	6775
Dp	6709	AGCGTGTGTATCCCACTGTCTGTATGTAGACGGGAAACAGGGCCAGGC---CAGATGGGC	6765
QY	6776	TCCAAGTTCAGGCAACACCGACCGCCAGGCTTCTTCAAGCACTGACCGAAGC	6835
Dp	6766	TCTAAGTCTCAGGCAACACAGCAGCGCCAGGCTTCTTCAATAGCTGAGAGC	6825
QY	6836	AACCTCCGCAATGTTCATAGTCCAGAAAGAAAGATGATCAAGAAAGCTGAAACCCCAAC	6895
Dp	6826	AACCTCCGCAATGTGTGATGTGAAAGAAAGAGAGATCAAGAAATCTCAACCCCAAC	6885
QY	6896	CGGATGAGCTGAAATCAATATCAGCCAGGCTGGGACGAGATCTTCAATATGCCGC	6955
Dp	6886	CGGAAAGAGCAAGAAATCAATATTGGCCAGGCTGGGAGGAAATCTTCAATGCCGC	6945
QY	6956	ATCACCGGAAACAGGCTTATATGACTATATGAACCAAGCGGTGTCAGAAATGTCAACAC	7015
Dp	6946	ATCACCTGAGAGAGGCTTATATGACTGTATGAAGCAAGGCGGTGTCAGAAACAGCCAGAC	7005

QY	7016	AAACAATGGGGGCTGGAGAGCGCAATATTGAAABAAGCACTCAATGGGTAAATATATGACACAGTGGGAA	7075
Db	7006	AAACAATGGGAGCTAAGAGCGCATTTATTGAAAGGACACTATGGGTAAATATATATGATCATGTGGGAA	7065
QY	7076	GAGTCCCCCGCGCGCTCAGCGCCAAATGCTTTTAACCTCTTGATGCGCAGTGGCAGCTGGCCC	7135
Db	7066	GAGCGCCCCCGCGCTCGCGCGCCAAATGCTTTTAACCTCTTGATGCGCAGCGCAGTGTGCC	7125
QY	7136	---GCTGCTATGCCCATTAACCGCTGTGACGGAACGAGTGAACAACACTCCTACTGTGCCA	7192
Db	7126	GCTGCTCTATGCGCCATAACCACTGTGACGGAACGAGTGAACAACGACTCACCCTCGCCA	7185
QY	7193	GGTGGCGGCGGGAAAGGCGCAAGTCTCTGGGAGAACCCAGCAGCCGGAABAAGCAAGTCCCGG	7252
Db	7186	GGTGGAGGTGGGAAABGCGCAAGGCTCTGGGAGAACCTAGCAGCCGGAABAAGCAAGTGTGCCA	7245
QY	7253	GCCCCGGGCGCTGGGCAATCTGGGAGAACCGGCGCAACCCCTGTCTCTCTCAGTGCACCTGGAGGGA	7312
Db	7246	GCACACAGGCTTACGCTGCCGAGAACGACCCCTTCTCTCTCTCAGTGCACCTCAGAGGGG	7305
QY	7313	GACTGCAAACCGCGGAGCGCGCTCAACCAACCGCGTGTGGAGGACAAGCCCTGTGTCCGA	7372
Db	7306	GACTGCAAATGCGCGAACAACACTCAACCAACGATGTGGAGGACCGGCGCTCATCTGCA	7365
QY	7373	GGTTCCACCGCATTTCCCTCTCAAAACCCCTGTATATAGGGGTGAGAGCGGGGTGTATAGGCT	7432
Db	7366	GGGTCCACCGCATTTCCCTCTCAAAACCTTTTATATAGGCTAATAGGATGATGTATATAGGCC	7425
QY	7433	TCCCAACCCCCACCGGCGCTCCCCGCGGGAGCGGGCGCCCTCGCTGGCCCCCAACAGCC	7492
Db	7426	TCCCGGCCCCCACTGGGCTTTGGGGACGGGAGCGGGCGCCCTTAGTGTGTCCCCACACAGCC	7485
QY	7493	TGGGACAGGAGAGCCCAAGCCACTGTCTGTCTGCAAGTACAGACACTTCTCCGACAGGAG	7552
Db	7486	TGGGATAGAGGAGCCCAAGCACTGTGTGTTCACAGTATGAGACACTCTCGACAGGAG	7545
QY	7553	TGACTCAAAACAGGGCGGGGGGGGGCGGGCGGTGEBUSPERLUTGACAGTTCACAGCAGC	7612
Db	7546	TGACCAAGGATTTGGGGGGGAGCGG-----TGCGAGGTCTCCGCAAA	7586
QY	7613	CACAGGAAAGGCGCTGTGAGGAGGGGGCGGCTGCGCACTCCCCAACCAAGGAAAGGAGCC	7672
Db	7587	GGCAGAAAGCAGCGCCAGCATGAGACAGACAGCTGTGACTCCGAGACTGAGAAAGAGCC	7646
QY	7673	CTTGAAGTCCGCGCTTGC-----GCCTCCATTCATCTGTCTCCGTCCAGAGCGG	7717
Db	7647	CTTGAAGTCTGCTGCTGCGCTCATCGCTCCGTCCGTCCATCTGTCTCATCTGATCAGAGGCTGG	7706
QY	7718	CATCTTGTGCTGTCTTAAAGCCTTAAGTCTACGATCTCCGCGCCCGGGCTTGGCCCTGTGAG--	7775
Db	7707	CATTC--TGCTGTCTTAAAGCCTTAAGTCTTAAAGTCTCCACCCCGGGCTTGGCCCTGTGAGTG	7765
QY	7776	--ACCTTACCTCAGGGGATGTTTACCTGTGCTGGGAGGAGGAGGAGGAA-----	7821
Db	7766	ACCTTACCTCAGGGGATGTTTACCTGTGCTCGAGGGGAGATGAGCAGAGGGGAGGG	7825
QY	7822	-----GGGGCGGGGAGGGGAGCAGCGCAGGCGTGTGGCAACCAACAGAGCGGCGCAGG	7876
Db	7826	ACAGCGGGCCAGAGAGGGGGGGGAGACGATCTGTGTGAGTCCGACTCTGTGCAATCTTGGGG	7885
QY	7877	CGGCGAGGAGCCCAAGACAGATGACCAAGCACCTTCACGCACTGTCT--CCCCGAAT	7934
Db	7866	TCAGCGGGAGCCCAACCAAGGCTGACCAAGGCACTTCATGCCACGCGCTGCGCCCTTAAC	7945
QY	7935	GCATTTGGAACCAAAAGTCTTAACTGAGCTGCGAGCCCGCGCGCTTCCCTCCGCTTCCCA	7994
Db	7946	GCATTTGGAACCAAAAGTCTTAACTGAACTCTCGGTGTCTTGGCCCTCTCTTCTGCCCCCA	8005
QY	7995	TCCCGGCTTAAGGCTCTGAGCAGATGAGCGAGGCGCTGTCAAGCGCCCAAGTGCAGCTGTT	8054
Db	8006	GCCCGGCTTGCCTTGCAGACAGAGCGTTCCAGCTTATCTTGCCTTAATGCTGTATCA	8065
QY	8055	CCGGTCTCCCAAGACTGCCCCAGCCCAACGATGTGCTGAAACCAAGTCAAGCGCGGTGG	8114

Db 8066 TCCAGTCTCCAAA---GGCCACCCAGCCCAAGACTGGAGGCCATGACAGCAGGTGG 8122  
Qy 8115 GCGGCAAAAAGGCGCAGAGTGGCGGCTGGGGGGAACGAGTCTCCAGCACTGAGTGT 8174  
Db 8123 GTGACACAAAGGGGCTGG-----CTGGGGACCGAGTCTTGCAAGAACTGACCGTT 8174  
Qy 8175 TTTTCAACATCGTTGCCGCGAGCGGTGGGAAGAAAGCAGATGTAAATGATGTGTGT 8234  
Db 8175 TCCGGGCTGTGTCTGTGGCAAGG---GAGGGAAGGACGATGAATG--GTGTTGGC 8228  
Qy 8235 TTTACAGGTATATTTTGTATCTTCAATTAATTAATTCAGATGTTTACGCAAGAAAG 8294  
Db 8229 TTAACAGGTATATTTTGTATCTTCAATTAATTAATTCAGATGTTTACGCAAGAAAG 8288  
Qy 8295 ACTTACCCAGTATTTACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8354  
Db 8289 ACTCGCCC-GTGTGTTTCTCCGCTGTGC-TTGGTCTCTACTTCTGTTTCAAGGCACT 8346  
Qy 8355 GCAGGCGCAGACGTGGTGAACCCCATCACTGCAAGCAAGGCGGCGGAGCTGTCTGTC 8414  
Db 8347 GC---CAGCAAGCGGTGGCCACCATACGCAAGCTTGGGGGTCAAGGGGCTCCGGCAC 8403  
Qy 8415 AGCGCCGCTGTCT 8474  
Db 8404 AGCGCACTGTG---CCCTTCCCAACCTTACCTTACGCAAAATGACCTGATCGTATTC 8460  
Qy 8475 TGTGGCGCATTTTGGCAGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8534  
Db 8461 TGTGGCGCTCTGTGCA---CGGCGGCACTTGTGTGTGTGTGTGTGTGTGTGTGTGT 8517  
Qy 8535 AAAAGCGAAT 8544  
Db 8518 AAAAGCAAT 8527

RESULT 7 8388 bp mRNA linear ROD 20-MAR-1999  
LOCUS AF113001  
DEFINITION Mus musculus silencing mediator of retinoic acid and thyroid hormone receptor alpha mRNA, complete cds.  
ACCESSION AF113001  
VERSION AF113001.1 GI:4454547  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
REFERENCE 1 (bases 1 to 8388)  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.  
ORDentlich, P., Downes, M., Xie, W., Genh, A., Spinner, N. B. and Evans, R. M.  
TITLE Unique forms of human and mouse nuclear receptor corepressor SMRT  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (6), 2639-2644 (1999)  
PUBMED 10077563  
REFERENCES 2 (bases 1 to 8388)  
AUTHORS Downes, M. R., Ordentlich, P. and Evans, R. M.  
TITLE Direct Submission  
JOURNAL Submitted (11-DEC-1998) Gene Expression Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, USA  
FEATURES  
source 1. 8388  
location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/tissue\_type="spleen; brain"  
626..8047  
/function="transcriptional co-repressor"  
/note="SMART alpha"  
/codon\_start=1  
/product="silencing mediator of retinoic acid and thyroid hormone receptor alpha"

/protein\_id="AAD20944.1"  
/db\_xref="GI:4454548"  
/translation="MSGSTQPVQOTRAAPRPYPHGISVYQJASHTDVGILLEYOH  
HREDYTHSLSPSIIOPORRRPSLSEPPGSESRSEBILPRESFFBELGPDLE  
TESKRLBELPDITLIRPSLATQPSSELTORSLAGLEPEVSPPHADBEL  
ELAPRLSKEBILQNDRIYREITVEQOISTLKKQQQLEBEAAPPREKVPSP  
IESKHSILQIYDENRKAABAHRLBELGQVLEPLYNQPSDTRQIENIKINAM  
RKULILYFRNRHARQWQRFQORDOLEMEKVEVERIENPRRAESKVEEYE  
KQPEIKQREIEMQSVQSGSLSGVSESEVEIIDGLSEONLEKQREYE  
AVIPMLYADDOARQIKFIMNGMLPMKYVORQVTMMSEBERPFEKQFQHK  
RGLJASFLERKTVARCVLYYLTKKNENKSLVRSRRRGSQQQQQQQQQQQ  
AASSQBEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEKKEK  
TANSQRRKRIITRSNANEETAPQSSLELSMKNESRMTVEEMKAKGL  
EHRWMSIARWVSKTWSQCKNFFYNGKRLNLEIIOQHKIKMEKKEANARKKKT  
PAASEETPPPALEDEMEASGANESELELEASQASNGEYVRECEGPAVN  
NSDPSVSPSPSEMKDTGPKPTGTEALPATOPVPVPEEPAVAPAPSPVDA  
PSPSPSHLPHPLMTMNMKRPLOLRPMRPSBLPRPMWKEPDEPSEEP  
PPEVSKSHKETEPEPEDEKAGTEALITVSAPLKEVBSKAATVTKSSGATODS  
DESATCSADEVDEPBGDKRLSPSPSLTAPAGDRASTSPQKPLDKQLKORAAI  
PEIOVTKVPEPREDTVPKVPVPVPPQIOLPBGDVSQSGSGRGRSPPVPAE  
KEAKPAPPAETGPKLPTPEPMSSGLPPIPREVITKSPHADSPASVYPPG  
HPLPLGHDSPAPVLPRLISNPPPLISAKHGVLEQLGALISQGMVOLRVPSG  
HAKAMPGLTMGLPLAVDPKLGTLGATSGSITKGLDSTRAADSPSRGSIITHTP  
ADVLTKGTSRITVGEDSPRLDRAEDTLPKGHVILEGKGVLTSEGMVSQSKKE  
DGRSSGPEHETAPAKRTYDMGEGVGRVTSASIEGLMGRALPQHSNLEQKHIL  
GSITQIPRSYVEAQEDYLRBAKLKEGTGPPPPPPPPDLTETKPRDLPLGPKL  
KPTHEGVAIVAEKRSIHEIPREELRTPELPLAPLKEGSIQGTPLKTDGAPL  
TQTKGVADRSIIGSPRPPLALPLDIMALDARACVRESLTPRLOBSLSKASIT  
RCAPVVPBELGKPROSPPLTEDHGAFTSHLDRGSPVTTREPTPLQBSLSKASIT  
DKLTVSTPEIKSHSTVPEHHPISYHHLRGVGTGVDLYRHGIIPLAPPTSLPR  
GIPLEAAAAAYLPHHLAPSPYPHLYPEYLRGVPDPLANRQIINDYVTSQOM  
HNAASAMORADMLRGLSPRESLALNVAAGRGILDSOVHPLPLVPEPGSTPA  
IDRLAVLTPAPPPSPSSRSLSPGSPHPLAKPTSPSSREBERREBERREBERRE  
TTVEHAPITRPTEDSSGAGSSSPASHTHNSPISPRQDLOORPSVLNHTSMKVG  
VTSVEGPTVLMARSTSSPVRPAATFPATPCPLAGTLEGYPLTMEVLVLPKE  
TSRVARPERAVDAGHAFITKPPGRPASSESPRSLAPSSHTAIATPAKPL  
APHASPPAPPTASADLREKTOGKPSIQLELSRLSYHAGAGSPGVPISPV  
SPSLTHDKLSKPLSELESKHLGELRHQPPGSLASAPALPHLRPLPEOSGSP  
LIQTAQKGHORVTLAONHSEVITQDTRHNPOLGSLPALPYSPGACPYVLD  
RRPSDILVLPREDHTPARGSPHSGKRSPPSPSTSVLSESDAIEIPSPBEMTER  
GHARSTAYPLTRDSEQSPKGLSPGTSOPPTFFSLTNSNANVSKINMCK  
LNTNHRNEBENYIQQPTEIFNMPPAITGAGLMTCSQAQVBSATNMGLAIIKRLM  
GYDQWSEPPPLGANAFNPLNLSAGLPAALAMITTDADGSDIALTSPGGGCAKYSGR  
PSSRKSPAPPLAGDRPPSVSVSHSEDCNRRTPLNVRVEDPSSGSPFPYNP  
LIMRLQAGYMASPPPPGLAAGSPLAGPHNANDEBPXPLCSQYTLDSSE"  
731..1387  
/note="Region: deleted in SMRT beta isoform encoded by  
Genbank Accession Number AF113002"

ORIGIN  
Query Match 58.2%; Score 4981.6; DB 9; Length 8388;  
Best Local Similarity 79.9%; Pred. No. 0;  
Matches 6351; Conservative 0; Mismatches 1265; Indels 331; Gaps 28;  
Qy 1 CATGTGGGCTCCACAGAGCTTGAGCAGACGTGAGAGGCGCACTGAGCCCGCTAACCC 60  
Db 625 CATGTGAGATCCACAGAGCTTGAGCAGACGTGAGAGGCGCTGCTGAGCCCGCTAACCC 684  
Qy 61 GCCCCAGAGCCCTTCTTACCCAGTCCAGATGCGCCCGAAGCAGACGAGCTGGGCTGCT 120  
Db 685 ACCCCATGGCATCTCTACCCGCGTCAAGTAGCCGCGTCCACAGAGAGTGGGCGTGTCT 744  
Qy 121 GAGATACAGACCACTCCCGGAGTATAGCCCTCCACACTGTGGCGGGGTCCATATCA 180  
Db 745 TGAGTACCAACACACACCCCGTACTACCTCAACCTGATACCCGGTTCATATCA 804  
Qy 181 GCCCAGCGCGAGAGGCGCTCCCTGCTGTGATGTTCCAGCCCGGGAATGAACGGTCCA 240  
Db 805 GCCACAGAGAGAGCGGCGCTCACTGTCTGATGATTCAGAGCTGGAGAGTGAACGTTCTCA 864  
Qy 241 GAGGCTCAACCTGGCGCAGAGTCCCACTATACCTGCCGAGGCTGGGGAATCAAGAT 300  
Db 865 GAGGCTCAACCTGGCGCAGAGTCCCACTATACCTGCCGAGGCTGGGGAATCAAGAT 924

QY	301	GGAGTCAATTGAAGCAAAGCCGCTTCCGCTTAGAGCTGCTCTGAACCCCTCTGTCGACG	360
Db	925	AGAAATTCACCCAGAGCAAAGCCGCTCGCTGAGCTACTACCCGATACCTCTGCTGCGCC	984
QY	361	GTCACCCCTGCTGGCGACCGGGCGAGCCCTGGGGGATCTGAACCTCAACAGGACCCGTAG	420
Db	985	ATCACCCCTGCTGGCGACTGGGGACCGAGTGGGCTTGAAGACCTTACAGAACCTTAG	1044
QY	421	CTTGACGGGCAAGCTGGAACCGATGTCTCCCCGAGCCGCCGCAACTGACCTGTAGCT	480
Db	1045	CCTGCGAGGCAAGGCTGGAGCCTGTGTGTCATCCCAAGTCCCCGCAAGCTGACCTGTAGCT	1104
QY	481	GGAGCTGTGTCGCCCACGGCTGTCTCCAGAGGAGCTGATCCAGAACATGACCGGTGGA	540
Db	1105	AGAGCTGGCGGCATCTCGACTGTCCAGAGGAGAGGTGATCCAGAACATGAGCCGTGGA	1164
QY	541	CCGAGAGATCAACATGGTAGAGCAGCAGATCTCTAAGCTGAAGAGAGAGCAGCAACGCT	600
Db	1165	CCGTGAGATCAACATGGTAGAGCAGAGATCTCCAGCTGAAGAGAGAGCAGCAACGCT	1224
QY	601	GGAGGAGAGGGCTGCCAAGCCGCCGAGCCTGAGAGGCCGTGTCAACGCCGCCATCGA	660
Db	1225	GGAAGAGAGGGCCGCCAAGCCGCCGAGCCGAGAACCGAGAGGCCGTGTGTCGACCAACCATGA	1284
QY	661	GTGCAAGCACCGGACGCTGTGTGCAATCATTTAGACGAGAACCCGGAAGAGGCTGAAGC	720
Db	1285	ATCAAGAGCACCGAAGCCTGTGTGCAATCATTTAGATAGAACCCGGAAGAGGCGGAAGC	1344
QY	721	TGCACATCGATTTGTGGAAGGCTGTGGGGCCCGAGGTGAGGCTGCGCTGTATCAACAGCC	780
Db	1345	CGCACACCGAATCTTAGAAGGCTGTGGGGCCCGAGGTGAGGCTGCTGTATCAACAGCC	1404
QY	781	CTCCGACACCCGCGAGTATCATGGAACATCAAAATTAACAGGCGATGCGGAAGAGCT	840
Db	1405	GTCTGACACAGCGCAGTATCAATGAATAATCAAAATTAACAGGCGATGCGGAAGAGCT	1464
QY	841	AATCTTGATCTTCAAGAGGAGGAATCAACGCTCGGAACAAAGGAAGCAAGTTCTGCA	900
Db	1465	GATCTTGATCTTAAAGCGAGGAACCAACGCGCAGAGAGTGGGAACAGCCTTCTGCA	1524
QY	901	GCGCTATGACAGAGTCAATGAGAGCCTTGGAAAAAGGTGAGCGCATCGAAAACAACC	960
Db	1525	GCGCTATGACAGAGTCAATGAGAGCCTTGGAAAAAGGTGAGCGCATCGAAAACAACC	1584
QY	961	GCGCGCGGGCCCAAGGAGCAAGTGTGCGAGTCTTACGAAAAAGCATTTCTTGAGAT	1020
Db	1585	GCGAAGAGGGCCCAAGGAGCAAGTGTGAGGAGTACTACGAAAAAGCATTTCTTGAGAT	1644
QY	1021	CCGCAAGCAGCGGAGCTGCAAGAGCGGATGCAAGAGCGAGGTGGGCCAGCGGGGCAAGTG	1080
Db	1645	CCGCAAGCAGCGGAGCTGCAAGAGCGGATGCAAGAGCGAGGTGGGCCAGCGGGGCAAGTG	1704
QY	1081	GCTGTCAATGTGGCGCGCCCGCAGCGAGCAAGAGGTGTACAGATCATCGATGGCCTTC	1140
Db	1705	GCTGTCAATGTGGCGCGCCCGCAGCGAGCAAGAGGTTTCTGAGATCATTTAGTGGCTTGC	1764
QY	1141	AGAGCAGGAGAACTTGAGGAAGCAAGTGTGCGCACTGGCCGTGATCCGCCCATGTGTGA	1200
Db	1765	TGAGCAGGAGAACTTGAGGAAGCAAGTGTGCGCACTGGCCGTGATCCGCCCATGTGTGA	1824
QY	1201	CGAGCGTGAACAGGAGGCGCATCAAGTTATCAACATGAAAGGGGCTTATAGGCGGACCCCAT	1266
Db	1825	CGAGCGGAGCAGGAGGAGATCAAGTTATCAACATGAAAGGAGCTCATGATGACCCCAT	1884
QY	1261	GAAAGGTGACAAAGCCGCGCAAGTCAATGAACATGTGAGTGAAGCAGAGAGAGACTT	1320
Db	1885	GAAAGGTGACAAAGCCGCTGAGTTTCCAAACATGTGAGGCGAGCAGAGAGAGACTT	1944
QY	1321	CCGGAGAAAGTTCAATGACAGCATCCCAAGAACTTTGGCCTGATCCGATATTTCTGAGAG	1380
Db	1945	CCGGAGAAAGTTTATGACAGCACTCTAAGAACTTTGGCCTGATTTGCTCATTTCTGAGAG	2004

QY	1381	GAAGACAGTGGCTGAGTGGCGTCCCTCTATTATCTACCTGACCTAAGAAGATTGAAGACTATTAA	1440
Db	2005	AAAGAAGCGGTGCTGAGTGTGTCTCTATTATCTACCTGACCTAAGAAGATTGAAGATTAAATTATCA	2064
QY	1441	GAGCTGTGTGAAGACGAGACTATCGGCGCCGCGCAAGAGCCAGCAGCAACAGCAGCA	1500
Db	2065	GAGCTTGTGTAGGCGAGGCTATGTGGCGCGGTGGDAGAGCCAGCAAGCAGCAGCAGCA	2124
QY	1501	GCACGACGACGACGACGACGACGACGCCCATGCTCCCGACGACGCCGAGAGAGAA	1560
Db	2125	ACAAACAGCAGCAGCAGCAGCAG-----ATGACGACGAGCAGCCAGGAGGAGAA	2172
QY	1561	AGATTGAAAGAGAGAAAGAAAGAGCGGAGAGAGAGAGAGAAAGCCGAGGTGAGAA	1620
Db	2173	GGAGGAAAGAAAGAAAGAGAGAGGCGCGCAAGAGAGAAAGAAAGCAGATCCGAGAA	2232
QY	1621	CGACAAAGAGAGACCTCCCTCAAGAGAGAAAGACAGACGACACTCAGAGGAGAGACAGCAG	1680
Db	2233	CGAGAGAGAAAGAACTCAGCAGAGAGAGAAAGACAGACGACCTTCTGGCGAGAGACACCATGA	2292
QY	1681	GAAAGAGGCTGTGGCTCCAAAGGCGGCAAACTGCGCAAGCCAGGAGAGACGCAAGG	1740
Db	2293	GAAAGAGGCGGTGGCTCCAAAGGCGCAAACTGCGCAAGCCAGGAGAGACGCGCAAGG	2352
QY	1741	CCGATATCCCGCTCATATGGCTTAATAGAGGCAACAGAGAGAGGCCATACCCCCAGCA	1800
Db	2353	CCGATATCCCGCTCCATGGCCCAACAGAGGCAACATAGAGAGACAGCCACCCCAAGCA	2412
QY	1801	GAGCGCCGAGCTGGCTCCCATGAGAGCTGAATGAGAGTTCTCGCTGACAGAGAAAGAAAT	1860
Db	2413	AAGTTCAAGACTGTGCTTCCATGAGATGAAAGAGAGTTCTCGCTGACTGAGAGAAAGAT	2472
QY	1861	GGAACAGCCAGAAAGGTCTCTTGAACACGGCCGCAACTGCTCGGCTATGCTCCGAGAT	1920
Db	2473	GGAGACAGCAAAAGAAAGGCTCTCCAGAAATGAGAGAGAACTGCTCAGCCTATCCGCGAT	2532
QY	1921	GGTGGGCTCCAAAGCTGTGTGCGAGGTGAATCTTCACTTCACTCAAGAAAGAGGCA	1980
Db	2533	GGTGGGCTCCAAAGCCGTGCTCCAGGTGAATCTTCACTTCACTCAAGAAAGAGGCA	2592
QY	1981	GAACTTGATGAGATCTTTCAGAGCAGCACAAGCTGAGAGATGAGAGAGAGAGAAAGCGCG	2040
Db	2593	GAACCTTGAAGAAATCTTTCAGAGCAGCACAAGCTGAAGATGAGAGAGAGAGAAAGCTCG	2652
QY	2041	GAGAGAAAGAAAGAAAGCGCGCGCGCGCGCAGCAGAGAGGCTGCATTCCCGCCGCTGT	2100
Db	2653	GAGGAGAAAGAAAGAAAGCCCACTGCGCGGAGCAGAGAGACAGCTTCCCACTGCGCG	2712
QY	2101	GGAGATGAGAGATGAGAGGCGCTGCGGCTGTAGCGGAAATGAGAGAGAGATGTGTGAGGA	2160
Db	2713	TGAGGACGAAAGATGTGAGACATCAGGCGCAATGTGCCAATGAGAAAGAGCTGTGCGAGGA	2772
QY	2161	GGCTGAAGCCTTTCATGCTCTGAGGATGAGGTGCCAGA---GGGAAATGCAGTGGGCC	2217
Db	2773	GGCAGAAAGCTCACAGGCTCTTGGAGATGAGTTTCCAGATTTGGGAGTGCATGTGCC	2832
QY	2218	AGCCACTGTCAACAACAGCTCAGACACCGAGACATCCCTCTCTCACACTGAGGCGCG	2277
Db	2833	AGCTGCTGTACAACAAGCTCTATATCTAGAGATGTGCCATCCCGGGTTCAAAAGCAT	2892
QY	2278	CAGAGACACAGGGCGAGATGGGCTCCAGGCTCCAGCCACCTTGGGCGCGCAGCGGACAC	2337
Db	2893	GAAAGACAC-----TGGGCTTAAACCACTGGCAGCTGAAGCATTGCCTGCTGCAC	2943
QY	2338	CCCAAGGCCACCCACCCCAACAGAGAGACATCCCGGAGCCCAATTGAGGCCACCCCGCG	2397
Db	2944	CCAGCCACTGTTCCTCTCTCAAGAGAACCGGAGATGAGCCCTGTGAGCTTCCCAT	3003
QY	2398	CTACTGAAGCCACCGAGGCCCTTACGCGCCCAACAGAGACCCCAATGGCTCTCTCAGCTCC	2457
Db	3004	CCCTGATGACAGTGG-----CCACACATCCCAAGAGCTTTCCTCATCACTGCGGACCCCG	3059
QY	2458	TCTCTGTGTTCCCAAGAGAGAGAGAGAGAGACCGCAGCAGCGCCCAAGTGAAGGA	2517

Db 3060 GGGTACTGTGAGCAAGGATGAAACAAGGCCCCGGCTGTCCAGCTCCCGACAGAGGA 3119  
QY 2518 GGGGGAGAGCAGAGCCCCCGGGCTAG-----GACCTGGCAG 2558  
Db 3120 TGCAGAGAGCAAGAGCTGAGGCCAGAGATCGATGTGGGAAAAACCAAGAGAGCCCG 3179  
QY 2559 TGGACACAGGGAGAGGCCAGAGACCCTCAAGACGAGTGCACGAGAGAGCCGAGAGG 2618  
Db 3180 AGGCTCTGAGAGAGCCCGGAGAGGTGTAAAGATGACCAAGAGAGAGACCGAGGAG 3239  
QY 2619 GGGCGGC---CAAGGCAAGAGACGCGAGAGCCGCTGAGGCAAGCGCGAGAGGAGCGCTCA 2675  
Db 3240 AGCTGAAGCAAAAGCCAAAGGCGACAGAGGCCATTGAACTGTGTGAGGACCACTTA 3299  
QY 2676 AGGCAAGAGAGAGAGGCGGAGCGGAGCGGACCACTGCGCAAGAGTCCGGCGCC 2735  
Db 3300 AG---GTGAGGAGGCTGTGAGCAAGGCAAGCTGTGACCAAGGATTCCAGCTCAGTGGCA 3356  
QY 2736 CCAGAGCAGCGACTCGAGTGTCTACCTGACAGTGCAGACGAGTGAATGAGCCGAGGAGCG 2795  
Db 3357 CCAGAGCAGTGAATCTTCACTGAGCACTTGACAGTGCAGATGAGTGAACCAACCGAAGAG 3416  
QY 2796 GCGACAAAGAACCGGCTGTGTCTCCCAAGAGCCGCTCTCAACCCGACTGGCGAACCCCG 2855  
Db 3417 GTGACAAAGGAGGAGGCTGTGTGTCAACAAGGCCAGGCTCTCAACCCGGGCTGAGATCCCG 3476  
QY 2856 GGGCCCAATGCTTCAACCCCAAGAGCCACTGGAAGCTGAAGAGCTGAAGCAGCGGCTG 2915  
Db 3477 GGGGCCAGTACTCTGCGCCAGAAAGCCGCTGGAAGCTGAAGAGCTGAAGCAGCGAGCGG 3536  
QY 2916 CCATCCCCCCCCATCCAGGTCAACCAAGTCCATGAGGCCCGCGGAGAGCGGAGCTCCCA 2975  
Db 3537 CCATCCCCCTTATCCAGGTCAACCAAGTCCATGAGGCCCGCGGAGAGACAGTACCCC 3596  
QY 2976 CCAAGCGAGCTCCCCAGGCCCAACCGCACACCGCAAAACCTGACCGGAGAGCGAGCGCC 3035  
Db 3597 CAAAGCGAGTTCCTCCCTGTGCTCCACCAACGAGAGCACTTACAGCGAGAGGTGAGTGT 3656  
QY 3036 CTCAGAGCTGTGAGCAGAGCCCCGGGGGCAAGAGCAGAGAGCCCGGACACCCCGCGGACA 3095  
Db 3657 CTCAGAGTGTGGAGGAGAGTCCACTGTGGCAAGTCCCGCAGCCAGTGTCTCTGCGGAGA 3716  
QY 3096 AGGA-----GGCTTCCGACGCCAGGCGCCAGAGCTGCTG 3131  
Db 3717 AAGAGGCAAGAAACCCGCAATCTTTCGGCTTTCACCACTAGAGGCCCAAAAGCTACCA 3776  
QY 3132 GGGACCCCCCTTGTGAGACTTCCGAGCTGACCTTCCCGTGCCCCCGTGAAGTATCA 3191  
Db 3777 CTGAGCCCCCAAGCTGTGATGAGGAGCTGCTTCCCATCTCTCAACGAGAGTATCA 3836  
QY 3192 AGGCTCTCCGCAATGCGCGAGACCTCTAGCTTCTCTAAGCTTCACCTGTGACCCAC 3251  
Db 3837 AAGCTTCCCCCAACGCGCGCTGAGCCCTTGTCTCTAACAACCCCCGGGTCAACCCG 3896  
QY 3252 TGGCTCTGGGCTCTCACTGACATGCGCGGCGCTCTGTGCGGCGCAACCACTATCTCA 3311  
Db 3897 TGCCTCTGGGCTCTCAAGATGAGTGCAGGCGCGCTCTGTGCAAG---TCCCCCATCTCA 3953  
QY 3312 ACCGCGTCCCCCTCATCTCTGTGCAAGACACCGAGCGTCTGAGAGGCAAAATAGTG 3371  
Db 3954 ACCCCCAACCTCATCTCTCTGTGCAAGATTCGGGCTTCTTGAAGAGGAGCTGGGTG 4013  
QY 3372 CCATCTCC---CAAGAAATGTGGGTCAAGCTCCAGCTCCGTACTCAGAGCATGCCAAG 3428  
Db 4014 CCATCTCCAGAGAGGGATGTCAAGTCAAGCTTGTGTGCTCACTCAGAGCATGCCAAG 4073  
QY 3429 CCGCGGTGGGCTCTGTCACTAGGGGCTGCTGCGCCATGAGACCCCAAAAGTGTGGAC 3488  
Db 4074 CCCCCATGGGCTCTCAACATGAGGAGCTGCTTTCGCTGAGACCTTAAGAAAGCTG--- 4129  
QY 3489 CCTTACGCGAGTGAAGCAGAGCAGTGTCCACGAGGCGCAGGCTGGGCGACCGAGA 3548

Db 4130 ----- 4129  
QY 3549 GCTTGGGGGTGCCACAGCCAGAGGCGCTCGTGTGAGAGGAGAGCTGTGGCTCAG 3608  
Db 4130 -----GGAGACGACTGGGCTCCG 4148  
QY 3609 TTCCGGGCGGAAGCATCAACCAAGGATTTCCAGCAACGCGGTGCGCTCGACAGCGCA 3668  
Db 4149 CACAGTGGAGCATCAACAGAGGCTCCCGAGTACCGGG-----CTGACAGACGGCC 4202  
QY 3669 TCACATACCGCGGCTCTCATCCACGAGCAAGCTGAGTACGTCTGTACAGGGCAACA 3728  
Db 4203 CCAGCTTACAGAGGCTTATCAACCAAGGACGCGCCCAAGCGTCTCTCAAGAGGTACCA 4262  
QY 3729 TCACAGAGTCAATCGGCGAGGACAGCCGAGTGTGAGACGCGGCGCGGAGGAGACGCC 3788  
Db 4263 TCAGCAGAGATCGTGGTGAAGACGCCAAGTGCCTTACCGGGGACAGAGAGACACC 4322  
QY 3789 TGCACAAAGGCGACGTTCATCTACGAAGGCAAGAGGCGCACGTCTTGTCTATGAGGGTG 3848  
Db 4323 TGCACAAAGGCGCATGTTCATCTATGAGGCGAAGAAAGCCAGCTCTATCTTATGAAGTG 4382  
QY 3849 GCATGTCTGTACCCAGTCTCCAGAGAGGACGCGCAAGACAGTCAAGACCCCCATG 3908  
Db 4383 GTATGTCCGTGTCAAGTGTCTTAAGAGAGATGAAAGAGCAGTCTGGGCCCAACCCATG 4442  
QY 3909 AAGCGGCGCGCCCAAGCGGCACTTATGACATGATGAGAGGCGCGGTGGGCAAGCCATCT 3968  
Db 4443 AAGCTGCGCCCTTAAACCACTATGACATGATGAGAGGCGCTGTGAAGGAGAGTGTCA 4502  
QY 3969 CCTCAGCCAGCATCGAAGGTCTCATGAGGCGGTGSCATTCGCGCGGAGGACACAGGCCCC 4028  
Db 4503 CTTAGCCAGCATGAGGAGCTCATGTGGCGCGCATCTCTGAGACGA-----CAGCC 4556  
QY 4029 ACCACTTAAAGACAGCAACATCCGCGGTCTATCAACAAGGATCTCTGTGCTCT 4088  
Db 4557 CCACCTCAAGAGGAGCATCATCCAGGCTCATACGCAAGGCACTCCGAGGTCTCT 4616  
QY 4089 AGGTGAAGGACAGAGAGACTACGTGCTGGGAGGCGCAAGCTCTTAAAGCGGAGGCA 4148  
Db 4617 ATGTGAAGGCGAGAGGACTTACTTACGGCGGAGGCGCAAGCTCTTGAAGCGAAGGGA 4676  
QY 4149 CGCTCCGCGCCCAACGCGCTCAAGGAGCCTGACCGAGGCTTACAGAGCGCAGGCGC---- 4204  
Db 4677 CACACACACCCCAACCAACCTCTGGGACCTTGAAGCTTACAGAGCTTACAGAGCTCCGCG 4736  
QY 4205 -----CTGGGCGCCCTGAAGCTGAAGCCGCGCCATGAGGCGCTGTGGCAAGTGAAG 4259  
Db 4737 ACCCTGTGGGTCCCTGAAGCTGAAGCGACCTCAAGAGGTTGTGTAGCACTGTGAAG 4796  
QY 4260 AAGCGGCGGCTTCATCATGAGATCCGCGGAGAGCTGCGGACACGCGCGAGCTGC 4319  
Db 4797 AGGCGGCGGCTTATTCATGAGATCCGAGAGAGGAGCTGCGCGCACACTGAGACTAC 4856  
QY 4320 CCTGCGCGCGGCGGCGCTCAAGAGAGGCTCATATCAAGAGGAGACCCGCTCAAGTAAAG 4379  
Db 4857 CCTGAGACCAAGGCTCTGAAGAGAGGTTCAATCAACCAAGGAGACCCCACTCAAGTAAAG 4916  
QY 4380 ACAACGCGCGCTTCACACTGTGCTCAAAAGACAGACGATAGCTCTCATTCGAGGCC 4439  
Db 4917 ACTGTGGGCAACCTTCATGTGGCAACAAAGAAACAGACGTTGGGTCTCATATTCGAGGCC 4976  
QY 4440 CCGGCGGAGGTTCCCAACCTGTGACCCGCTGAGATGTATGCGCGAGCGCCGGGCACTGG 4499  
Db 4977 CCGGCGGCGCTTTCCTGTGCGACCCGCTGACATTAATGTGCTGACGCGCGGCACTGG 5036  
QY 4500 AAGGTGCTGTACAGAGAGAGCTGAAGAGCGGCAAGGACCGGACAGCTCGGGGG 4559  
Db 5037 AAGGTGCTGTATGAAGAGTGTGAAGAGCGGCAAGAGCCAGACATGTGTGACAGGG 5096  
QY 4560 GCTTCATTGCGCGCGCGCGCTGATCTTGTGCTTGAAGTGTGATAGCCGCGAGAGGCC 4619  
Db 5097 GCTTCATCAACAGTGGGCGCTCAAGTGTGTGCTTGAAGTGTGGCAAGGACGCGAAAGGCC 5156

QY 4620 CCTGACCTATGAGGACCAAGGGGACACCTTTGCGGCGACCTCCACAGAGGTGCGCG 4679  
DB 5157 CACTGACTTTGAAAGAACCAAGGGGACCTTTGACCAAGTCACTGGCAGGTGGCTCCCG 5216  
QY 4680 TGAACATGCGGAGCGCCACGCGCGCTGACAGAGGGACGCTTTCTGTCAGCAAGGCAT 4739  
DB 5217 TGAACCAAGAGGAGCCACGCGCATCTTCAAGAAAGGACCTCTTATCCAGCAAGGCGT 5276  
QY 4740 CCCAGAGCGGAAAGGAGTGAAGCGCTGTAAGATGCGCAAGTCCCGCAGAGACCG 4799  
DB 5277 CCCAGAGCGGAGGCTGACATTTACACCCGGAGATGCGCAAGTCCCAAGACACTG 5336  
QY 4800 TGCAGAGACCAACCAACACCCCATCTGCGCTTATGAGACACCTGCTTGGGGCTGAGTG 4859  
DB 5337 TGCCGAGAGACACCTTCAACCCCATCTGCTTATGAGACCTTGTCCGGGGCGTGACTG 5396  
QY 4860 GCGTGAACCTGTATGCGACGCAATCCCTGAGCTTGAACCCCACTTCATACCCCGG 4919  
DB 5397 GTGTGAACCTGTACGTTGGTCAATCCCATTTGGCCTTTGACCCCACTTCATACCCCGAG 5456  
QY 4920 GATCCCTCTGGA---CGAGCGCGTGCCTACTACCTGCGCGGACACCTGGCGCCCAAC 4976  
DB 5457 GGATCCCTTGAAGAGAGAGCGCGACGCTTACTGCTGCGCGGACCTTGGCGCCCAAGC 5516  
QY 4977 CCACCTACCCGACCTGTACCAACCTTACTCATCCGCGCTAACCCGACAGCGGCGCG 5036  
DB 5517 CCACCTACCCCAACCTGTACCAACCTTACTCATCCGCGGCTAACCTGACAGGGGCGCC 5576  
QY 5037 TGAAGAACCGGAGACCATCATCATATGACTATCATCATCTTGCAGAGATGACACACAA 5096  
DB 5577 TGAAGAACCGGAGACCATCATCATATGACTATCATCATCTTGCAGAGATGACACAAAG 5636  
QY 5097 CGGCGACCGGACATGCGCGGAGCGTGAATGTGAGGGGCGCTGCGCGCGCGAGTCT 5156  
DB 5637 CTGCTCGGCGACATGCGCGGAGCGTGTGACATGTGAGGGGCTGTGACCGCAGAGTCT 5696  
QY 5157 CGCTGGCACTCAACTACGCTGCGGGTCCCGGAGGCATCATGACCTGTCCAAAGTGCAAC 5216  
DB 5697 CGCTGGCGCTTAATTAATGCGCGTGGCGCCAAAGGCAATTAATGACCTGTGCCAAGTGCAC 5756  
QY 5217 ACTTCGCTGTGCTCGTGGCGCGGACCAAGGACACCCGACCAACCGCATGAGACCGCTTG 5276  
DB 5757 ACTTCGCGCTGTGCTGGCGCGGACCAAGGACACCCGCTCAACCGCATGAGACCGCTTG 5816  
QY 5277 CCTTACTCCCAACCGGCGCGGACCGCTTGAAGACCGGCGGACAGCTCCCACTCTCCG 5336  
DB 5817 CCTTACTCCCACTGTGCGCGGACCGCTTGAAGACCGGCGGACAGTACGCTGCTGCC 5876  
QY 5337 CAGAGGTCCCAACACTTGAAGAAACAAACCAACGATCTGTCGAGCGGAGCGAG 5396  
DB 5877 CAGAGGCGCCCACTCACTGATTAACCAATGCGCACTTTCATGAGAGCGG----- 5929  
QY 5397 ACCGGGATCGAGAGCGGAGACCGGAGTCGGAGCGGGAAAGTCCATCTCACGTCACCA 5456  
DB 5930 -----GAAAGGAAACGTGAGCGGAAAGAAAGTCAATCTCAAGTCAACA 5978  
QY 5457 CGACGCTGAGACAGCAACCACTTGTGAAGCTGTGTACAGAGACAGGCGGACAGAGG 5516  
DB 5979 CTACAGTGAAGATGACCACTGTGAGACCTGTGTACGAGACAGACAGCGGGGCTGGGG 6038  
QY 5517 GAGAGAGCGGCGGGGGGTGGGGGAGAGAGAGCGCGCGCTGCCATGCCATGCGACCG 5576  
DB 6039 GAGAGAGCGGCG-----CGGCTTCCACATCCAC 6068  
QY 5577 AGCACTCGGCCACTTCCCTCGGACCCAGAGATGCTTCAGAGACAGAGCCCAAGTGTCTTC 5636  
DB 6069 AGCACTCGGCCACTTCCCTCGGACCCAGAGATGCTTCAGAGACAGAGCCCAAGTGTCTTC 6128  
QY 5637 ACACACAGAGCATGAAAGGTATCATACGCTGTGAGAGCCAGCAAGCCCAAGGTCTCTG- 5695  
DB 6129 ACACACAGAGCATGAAAGGTGTGATCACTCGTGAAGCCCGCAGCGCGCTGCTGA 6188

QY 5696 -----AGGTCAACTTCAACCTTCTCAACCCGTTGCGCCAGCTGCGCACTTCCACTG 5747  
DB 6189 GATGGGCGAGGTCACCTTCCACTTCTGCGCTGTCCGCGCAGCTGCGCACTTCCACTG 6248  
QY 5748 CCACCACTGCGCACTGGGCGGACCCGTCGATGAGGGGTCTACCTTCACTGAGAGCCG 5807  
DB 6249 CCACCACTGCGCACTTGTGACACTTGAAGGGGTCTTACCTTCACTGAGAGCCG 6308  
QY 5808 TCTTGTGCGCCAGAGAGGCGCGCGGGTGCCTGCGCGGACAGAGCGCGCGGACAGCACCG 5867  
DB 6309 TCTTGTGCGCCAGAGAGACTTGTGAGTGCCTGCGGCTGCGCGGCGCGGAGCGGCTG 6368  
QY 5868 GCATGCTCTTCTGCGCAAGCGCCGCGCTCGGGCTGAGAGCGCGCTTCTCCCA 5927  
DB 6369 GCATGCTCTTCTTCAAAACCCCGGCGG-----GAGCGCGCTTCTTACCA 6419  
QY 5928 GCAAGGGCTCGAGAGCGCGCGCGCTGATGCTGCTGTCGCGGACGCGCATCGCGC 5987  
DB 6420 GCAAGAGCTCGAGAGCGCGGATGCTGAGCAACCCGAGCTTCAAGCAACAGCATGCGC 6479  
QY 5988 GCACCTGCGAAGAACTCGACCTTCAACGCGCAGCGCGGACCGCGCGCGCACTG 6047  
DB 6480 GCACCCGCAAGAAAGAACTTGAACCCCAACATGCGAGTCCGAGCCCGCGCGG---CCCA 6536  
QY 6048 CCGGCTTGGAGCCCGCAACCGGAAAGACTCAAGGTAAACCTTTTCACTCCAGAAC 6107  
DB 6537 CCTCGGCTTCAAGATGCAACGAGAAAGAACTCAAGGTAAACCTTTTCACTCCAGAAAT 6596  
QY 6108 TGAACCTCGGTTCTTGGGTTTACAC---GGCAGAGCTACAGCGCGGAGGGGTGAGC 6164  
DB 6597 TGAACCTCGGTTCTTGGGTTTACACAGTGAAGCTGAGCTTACAGCGCGCGGAGGGGTGAGC 6656  
QY 6165 CCGTGAAGCTTGTGAGCTTCAACCAAGTGAACCAAGCAAGAGGGGTCTCCCAAGCACTG 6224  
DB 6657 CCATGAGCGCGGTGAGCTTCCCGAGCTGACCCAGCAAGAGAGGGGTCTTCCAAACCTGCG 6716  
QY 6225 AAGAGCTGACAGAGGCACTGAGGGGAGCTGCGCGCCCAAGCAAGCGCGCGCTGA 6284  
DB 6717 AAGAGCTGAGAGAGGCACTTGAAGGGGAGCTGCGGACAAAGCAGCGCGCGCATAGA 6776  
QY 6285 AGCTTGGGAGGAGCGCGCGCGCACTTCCCAACCTGCGCGCGCTGCTGAGAGCGAGCGCT 6344  
DB 6777 AGCTTGGGAGGAGCGCGCGCGCACTTCCCAACCTGCGCGCGCTGCTGAGAGCGAGCGCT 6836  
QY 6345 CGTTCAGCGCGCTGTCAGAGCGCGCGGAGGTCAAGGTCAACGAGCGGCTTCAACC 6404  
DB 6837 CATTCAGCGCGCTTCCAGAGCTGCGCGGATCAAGGTCAACGAGGGGTGATACCC 6896  
QY 6405 TGGCCAGACATCATGAGGTGATCAACAGAGCTACACCCGCGCAACCCCAAGCAGAC 6464  
DB 6897 TGGCTCAAGACATCAAGCGAGGTCAATTAAGAGAGCTTACAGCGCGCAGCAGCAGC 6566  
QY 6465 TGAAGGCAACCTTGGCGCGCGCGCTTACTCTTCCCTGGGGGCGAGCTGCGCGCTG 6524  
DB 6957 TCAAGTGGCGCGCGCGCGCTTCTTACTCTTCCCGGAGCGAGCTGCTGTCTGG 7016  
QY 6525 ACTTCGCGCGCGCGCAACCAAGTGAACCTTCACTTCCCGCGCGCGGACCATGATGCGCGCG 6584  
DB 7017 ATCTTGGCGCGCGCGCAACCAAGTGAACCTTCACTTCCCGCGCGCGGACCATGAGCGCGCG 7076  
QY 6585 GTGCTCCCGCGCGCAACGAGAGGAGGCAAGAGTTCAGAGGCGCAACAGAGTGTCT 6644  
DB 7077 GGGGATCCCGCGCAAGTGAAGGGGCAAAAGGTCCCGAGAACCGAGCAAAACATCGGTCC 7136  
QY 6645 TGGGTGTGTGAGGACGATATTGAACCTGTGTCCCAACCGAGGGGCAATGACGAGCGAG 6704  
DB 7137 TGGGCAAGAGCGAGGATGCAATGAGCTGTGTCCCAACAGAGGGCAATGATGAGCGAG 7196  
QY 6705 GGCACCTCCCGAGTGTGTGATCCCGCTGCTGTACCGGGATGGGGAAAGAGAGCGCA 6764  
DB 7197 GACATGCTTGGAGCACTGTGTACCACTGTGTATCGAGACGGGAAACGGGCGAGCG--- 7254  
QY 6765 GCAGGATGGGCTCAAGTCTCCAGGCAACCAAGCGAGCGCGCAAGCTTCTTCAAGCAAGC 6824

Db 7255 -CAGAGTGGGCTAGAGTCTCCAGGCAACACGACGAGCCGCAACCTTCTTCACTAGTAAG 7313  
Qy 6825 TGACCGAGAGAACTCCGCCATGGTCAAGTCCAGAGAGAGAGATCAACAAGAACTGA 6884  
Db 7314 TGACTGAGAGAACTCCGCCATGGTGAAGTGAAGAAAGCAGAGATCAACAAGAACTGA 7373  
Qy 6885 ACACCCCAACCCGAGATGAGGCTGAATATCAATATCAGCAGCCTGGAGCGAGATCTTCA 6944  
Db 7374 ACACCCCAACCCGAGAGAGAGAGAGATCAATATGAGCAGCCTGGAGCGAGATCTTCA 7433  
Qy 6945 ATATGCCCGCATCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7004  
Db 7434 ACATGCCCGCATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7493  
Qy 7005 ATGACAGACCAACAATGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7064  
Db 7494 ACAGCAGACCAACAATGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7553  
Qy 7065 ACCAGTGGAG 7124  
Db 7554 ATCAGTGGAG 7613  
Qy 7125 CCAAGCTGGCC--GCTGCTATGCGCAATACCGCTGCTGACGAGAGAGAGAGAGAGAG 7181  
Db 7614 CCAAGTGGCCCGCTGCTGCTATGCGCAATACCACTGCTGACGAGAGAGAGAGAGAGAG 7673  
Qy 7182 TCACCTGGCCAGAGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7241  
Db 7674 TCACCTGGCCAGAGTGGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7733  
Qy 7242 CCAAGTGGCCCGGCGGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7301  
Db 7734 CCAAGTGGCCAG 7793  
Qy 7302 ACTCGAG 7361  
Db 7794 ACTCGAG 7853  
Qy 7362 CTTGCTGGCCAGAGTTCACAGCAGATTCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 7421  
Db 7854 CTTGCTGGCCAGAGTTCACAGCAGATTCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAG 7913  
Qy 7422 GTGCTAGAGTTCCTCCACAG 7481  
Db 7914 GTGCTAGAGTTCCTCCACAG 7973  
Qy 7482 CCCACAG 7541  
Db 7974 CCCACAG 8033  
Qy 7542 CCCACAG 7601  
Db 8034 CCCACAG 8074  
Qy 7602 TCCAG 7661  
Db 8075 GTCCCGACAG 8134  
Qy 7662 AG 7706  
Db 8135 AG 8194  
Qy 7707 TCCAG 7766  
Db 8195 TCCAG 8253  
Qy 7767 CCGTGTGAG--ACCTTACTCAGAGAGAG--GTTTACTGTGCTCGGAGAGAGAGAGAGAGAGAG 7820  
Db 8254 CCGTGTGAG 8313  
Qy 7821 AGGAG 7847

Db 8314 GACAG 8340  
RESULT 8  
AF113002 7465 bp mRNA linear ROD 20-MAR-1999  
LOCUS Mus musculus silencing mediator of retinoic acid and thyroid  
DEFINITION hormone receptor beta mRNA, complete cds.  
ACCESSION AF113002  
VERSION AF113002.1 GI:4454549  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Muridae; Murine; Mus.  
REFERENCE  
1 (bases 1 to 7465)  
Odentlich,P., Downes,M., Xie,W., Genin,A., Spinner,N.B. and  
Evans,R.M.  
Unique forms of human and mouse nuclear receptor corepressor SMRT  
Proc. Natl. Acad. Sci. U.S.A. 96 (6), 2639-2644 (1999)  
2 (bases 1 to 7465)  
Downes,M.R., Odentlich,P. and Evans,R.M.  
Direct Submission  
Submitted (11-DEC-1998) Gene Expression Laboratory, The Salk  
Institute for Biological Studies, 10010 North Torrey Pines Road, La  
Jolla, CA 92037, USA  
FEATURES  
source  
1..7465  
Location/Qualifiers  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/tissue\_type="spleen; brain"  
363..7124  
/function="transcriptional co-repressor"  
/note="mSMRT beta; longer SMRT alpha isoform encoded by  
GenBank Accession Number AF113001"  
/codon\_start=1  
/product="silencing mediator of retinoic acid and thyroid  
hormone receptor beta"  
/protein\_id="AA020945.1"  
/db\_xref="GI:4454550"  
/translation="MSGSTQPVAVQWRAEPRYPHGISYVQIARSHPTPLNPSDT  
ROYHENTKINQAMRKXILYFKRBNHAKMORFQCORYQOMAEWKYERITENNR  
BRAKESKYRREYREKQPEIRKQRELOMRNSQGRSGISMABSEHYSYETIDIL  
SEBENLEKQROLAVIPMLYDADQQIKITNNMGHDDPKYTKDQVTHWMEQER  
DTRERKQHPKNGFLIASFLERKTVAECVLYYLTKKNYKSLVRSYRRKSGQO  
QOQOQOQOQOMARSSQOEKEKEKEADKEKEDANKEKELSKETDIDTSGBD  
NDEKAAVASGRKTANSQGRKGRITRSMANENHETATPQOSSLASMEMNNSHW  
TEERMETAKKGLIHHGNMSAIRMVSKTVSCKNKNFYVYKRONIDILQHKLM  
EKEBNARRKKKTPALASEBETAPPADEMEASGASNBEELAEABASQASGNY  
PRVGECSGPAVANNSSDTEVSPSRISATYDQTKPTGTALPATOPVPPPEEPAV  
APABSPVPDASGPSPSPESHNLPHPLTLTRNKKRRLQLPRQMRRSRLPRS  
MMKPEPEEASERKPKSKSDHKETEEBEDKALGETALIEVS EADLVKEAGSKAA  
VTGSEBEGATODSSATCSADYDEDEGDKRLSPRSILTPADPASTPQF  
LDLKKQRAAALPIVYTKYHRRPRLPKYPPVPPYPOHLOPBGDVSQGGSP  
RGSRSBPVPAEKAEKPAFPAPFTBGPPLTPRPWSSGAPPIPREVITKTSPPHA  
ADPSAFSTPPGHPPLGLHDSARPVLPRLISNPPPLISAKHGLEROLGAI50Q  
GMSVGLRVPHSEHAKPMGLTWGLPLAVDPKGLGTALGATSGSITKGLPSTAAAG  
PSYRGSITHTGPADVLKGTISRIVGDSRSLDRABEDTLPKGHVIEGKSHVLSY  
EGMSVSGCSQEKEDGSSGPHETAPAKRYDDMSRGLVGTSTVASIEGLMGRAPRO  
HSHLAKQHRIKRSITOGIPRSYVEADQEDILRABAKLKRGGPPPPPPRDLLETYK  
PRULDPGLKLPKTHGEGVATYKEAGRSIHETPRELRTPEPLAPRPLKBSITQ  
GTPDKTDSGAPSTGKHDVRSITGSGRPPLALPDIDADALRACCTEBSLKR  
SGTSAGSGSITRGAAPVVPBLKPRQSPPLYEDHGNPFTSHLPRGSPVTRREBPL  
QESGLSSKASQDKLSTPRELAKSSTVPEHNPAPISYBHLHAGVGLDLYRHG  
IPLAFDPTSIPRGLPLEAAAAYLLPHLAPSPYPLPYPLIRGYPDPAALNRQT  
IINDYVTSQOMHNAASAMAORADMLRGLSPRESSIALANYAAGPRGIIDISOVPLV  
LVPTPTPTATLDRALVLETPAPRPSRHSSTPLSGGTHLAKTPATSSSBERER  
ERERKDSILSTTIVETVAHPIWRPTESSSAGSASRASHTHQSHSPISPTQDLQOR  
PSVLAHNTSMKGVNTSVEPGTPVLRMARSTSSPVPAPATFPATHCPLAGTLEGV  
PTLMEPVLTPKETSRVAPERRARVDAGHAFLTLPGRPREPASPSKSSBPSRLAPSSS





Dh 2349 GCCATTGAACGTGTCTGAGGACCACTTAAG---GTGGAAGAGGTGTGAGCAAGGCA 2405  
Qy 2705 AGGACCACCACTGCGCAAGAGCTCGGCGC-----CCCGAGAGCAGAGACTCGAGTGACTGC 2764  
Db 2406 GCTGTGACCAAGGGTTCAGAGTTCAGAGTCCACCAAGACAGTGAATCCAGTGCACCTGC 2465  
Qy 2765 AGTCAAGACGAGGTGATGAGCGGAGGCGGCGACAAAGAACCGGCTGTCTGTCCCAAG 2824  
Db 2466 AGTCCCATGATGATGAGCGAACCAGAGAGGTGACAAAGGCGAGCTGTCTGTACCAAG 2525  
Qy 2825 CCCAGCCCTCTCAACCCGAGCTGAGGACCCCGGCGCAATGCTCAACCCCAAGACCACTG 2884  
Db 2526 CCCAGCCCTCTCAACCCGAGCTGAGATCCCGGCGCAATGCTCAACCCCAAGACCGCTG 2585  
Qy 2885 GACCTGAAGCAGTGAAGCAGCAGCGGCTGCATCCCGCCATTCAGAGTCAACAAAGTC 2944  
Db 2586 GACCTGAAGCAGTGAAGCAGCAGCGGCTGCATCCCGCCATTC-----GTCAACAAAGTTC 2642  
Qy 2945 CATGAGCCCCCGGAGAGACGAGCTCCACCAAGCCAGCTCCCGCAGCCCAACCGGCA 3004  
Db 2643 CATGAGCCCCCGGAGAGACAGATACCCCAAGCCAGTTCCTCCGTGTCTCAACCC 2702  
Qy 3005 CCGCAAAACCTGACCGCGAGAGAGAGCGCCCTCAAGAGCTGTGAGAGAGCGCCCGGAGC 3064  
Db 2703 ACGCAGACCTTAACAGCCAGAGGATGACGTGTCTCAGCAGTGTGGAGAAATCCACGTGGC 2762  
Qy 3065 AAGACAGAGAGCCCGGACCCCGCGCGACAAAGG-----G 3100  
Db 2763 AAGTCCCGAGAGCCAGTGTCTCTGCGAAGAAAGGAGCAAGAAACCGCATTCCTTTCG 2822  
Qy 3101 GCCTTTCAGCGGAGGCGCCAGAGAGCTGCTGGGAGCCCGCTTGTGACTTCCGAGCTG 3160  
Db 2823 GCTTTCCCACTGAGGGGCCAAAGCTAACGACTAGAGCCCGCAGCCGCTGTATCGGGCTG 2882  
Qy 3161 CCTTTCCTGCTGCGGCGGCGGCGGAGTGAATCAAGGCTTCGCGCAGTGCCTGGAGCCCTCA 3220  
Db 2883 CCTTTCCTGCTGCGGCGGAGGATCAAGACTTCCCAACGCGCGCTGAGACCCCTCT 2942  
Qy 3221 GCCTTTCCTAAGCTCAACCTGATCCACCTGCGGCGCTGAGACCTGAGCGCGG 3280  
Db 2943 GCTTTCCTCAACACCCCGGATCCCGGCTGCTGAGGCGCTGCACAGATGTGCCGG 3002  
Qy 3281 CCGGTCCTGCGCGCGCCACCCACCATCTCCAAACCGGCTCCCTCATCTCTCGCCAG 3340  
Db 3003 CCGGTCCTGCGAGG---TCCCCCATCTCTAACCCCCACCCCTCATCTCTCGCCAG 3059  
Qy 3341 CACCCAGCGCTCTGAGAGGCAAAATAGTGCATCTCC---CAAGAAATGTGCTCAG 3397  
Db 3060 CATCCCGGCGTACTTGAGAGGCAAGCTGGTGCATCTCCAGCAGGGAGTGTCAGTCCAG 3119  
Qy 3398 CTCACGCTCCCTAATCTCAAGACATGCAAGGCGCCCGGTGGGCTCTGTCACTATGGGCTG 3457  
Db 3120 CTTCTGTGTGCTCACTCAAGACATGCAAGGCGCCCATGGGCTCTCACTATGGGCTG 3179  
Qy 3458 CCGCTGCGCAATGAGCCCAAAAAGCTGAGCACTTCAAGCGAGTGAAGCAGAGCAGCTG 3517  
Db 3180 CCGCTTGCCTGTGACCTTAAGAAAGCTG----- 3206  
Qy 3518 TCCCAAGGAGGCGCAGCTGAGGCGCACCGAGAGCCTGAGGAGTCCACAGGCCAGAGGCG 3577  
Db 3207 ----- 3206  
Qy 3578 TCCGTGTGAGAGGAGACAGCTTGTGGCTCAGTTCCGAGCGGAAGATCAACCAAGGCAAT 3637  
Db 3207 -----GAGACAGACTGGGCTCGGCCACAGTGGAGACATCAACCAAGGCGCTC 3254  
Qy 3638 CCGAGCAGCAGGAGGCTCTGAGCAGAGCCATCAATACCGGCTCATCAACCAAGGCG 3697  
Db 3255 CCGAGTACCGGG-----CTGACAGAGGCGCCCGCTCAAGAGGCTTATCAACCAAGGCG 3308  
Qy 3698 ACGCAGACTGACGTCTGTATCAAGAGGACCAATCAACAGATCATCGGCGAGAGCAGCGC 3757  
Db 3309 ACGCCGCGAGAGCTCTCTACAGAGGATACATCAAGAGATCGTGGTGAAGAGCAGGCCA 3368

Qy 3758 AGTGTCTTGAACCGCGCGCGGAGAGACAGCTGTGCCAAAGGCGCACGTTCATCTAAGAGGC 3817  
Db 3369 AGTGTCTTGAACCGGAGCAGAGAGACAGCTGTGCCAAAGGCGCACATTCATCTAAGAGGC 3428  
Qy 3818 AAGAAAGGCGCACGTCTGTCTTATGAGGATGAGTGTCTGTAGCCAGTGTCCAAAGAG 3877  
Db 3429 AAGAAAGGCGCACGTCTTATCTTATGAAGTGTGTGTCTGTGTCAAGTGTCTTAAGAG 3488  
Qy 3878 GAGGCGAAGAGACTGAGAGACCCCGCATGAGAGCGCGCGCGCCCAAGCGGACCTATGAC 3937  
Db 3489 GATGAAGAGAGCAGCTGAGGCGCACCCCATGAGACTGCGCGCCCTTAAGCAGCATATGAC 3548  
Qy 3938 ATGATGAGAGGCGCGCTGTGGGAGAGACCATCTCTCAGCCAGCATGAGAGTCTATGAGC 3997  
Db 3549 ATGATGAGAGGCGCGCTGTAGGAGAGACTGTACCTCAACAGCATAGAGAGCTATGGGC 3608  
Qy 3998 GTGTCATTCCTGCGGAGGAGCAGAGCCCGCACCTCAAAAGAGCAGCACCAATCCG 4057  
Db 3609 CGGCGCATCCCTGTGAGCAGC-----ACAGCCCGACCTCAAGAGAGCAGCATCATCCGA 3662  
Qy 4058 GGTTCATTCACAGAGGATTCCTCGGTCTTACGTGAGAGGACAGAGAGTACTGTGCT 4117  
Db 3663 GGTTCATTCACAGAGGATTCCTGAGGTCTTATGTGAGAGGCGAGAGACTACTTACG 3722  
Qy 4118 CGGAGGCGCAAGCTCTTAAGCGGAGGAGGACGCGCTCGCGCCCGCCTCAAGGAGC 4177  
Db 3723 CGGAGGCGCAAGCTCTTGAAGCGAGAGAGACACACACCCCGACACCACTCGGAGC 3782  
Qy 4178 CTGACCGAGGCTTAAGAGAGCAGGCG-----CTGGGCGCCCTGAAGCTGAAGCG 4228  
Db 3783 CTGACTGAGACTTAAGAGCCCGGCGCTGTGAGCCCTGTGGGTCTCTGAAGCTGAAGCG 3842  
Qy 4229 GCCATAGAGGCGCTGTGTGCGCACGAGTGAAGAGGCGGCGCTCATTCATGAGATCCG 4288  
Db 3843 ACTCACAGAGGTGTGTGAGAACTGTGAAGAGGCGGCGCTTATTCATGAGATCCG 3902  
Qy 4289 CGCAGGAGCTGCGGCGCACAGCCCGAGCTGCGCTGCGCGCGGCGCTCAAGAGAGGC 4348  
Db 3903 AGAGAGAGCTGCGCGCGCACACTGTAGCTACCTTGGACACAGGCGCTGTGAAGAGAGGT 3962  
Qy 4349 TCCATCAACGAGGAGCACCGCGCTCAAGTACAGACCGGCGGTTCACACTGTGCTCAAA 4408  
Db 3963 TCCATCAACGAGGAGCACCCACTCAAGTACAGCTGTGGGAGCCCTCATGTGGACCAAG 4022  
Qy 4409 AAGCAGCAGTATGCTCTCTCATATGAGAGCCCGGCGGAGAGCTTCCACCCGTGACCCG 4468  
Db 4023 AAACACGAGCTGCGCTCATCATATCGGAGCCCGGCGGCGCTTTCCTGCGCTGACCCG 4082  
Qy 4469 CTGAGTGTGATGAGCGGCGCGCGCGGCGCATGGAAGTGTGCTTACAGAGAGGCTGAG 4528  
Db 4083 CTGAGCATATGTGTGAGCGCGCGGCGCATGAGGTGTCTGTATGAAGAGAGCTGAG 4142  
Qy 4529 AGCGGCGCAGAGAGCCCGCAGCAGCTCGGAGGAGCTCATTTGCGGCGGCGCGCGCTCAT 4588  
Db 4143 AGCGGCTCAGAGAGCAGAGATGTGAGAGGAGGCTCATACAGTGGGCTCCAGTGTCTC 4282  
Qy 4589 GTGCTGAGCTGTGATAGCGCGGAGAGGCGCGCTGACTATGAGAGCAGCGGAGCACCC 4648  
Db 4203 GTGCTGAGCTGTGAGAGGAGCAGCGCAAGGCCCACTGACTTAAGAGAGCAGCGGAGCACCC 4262  
Qy 4649 TTTGCGGCGCACCTCCACAGAGTGTGCGCGGTACCATGAGGAGCCCAAGCGCGGCTG 4708  
Db 4263 TTCAACAGTACCTGCGAGTGTCTCTGTGTACCAAGAGGAGCCCAAGCGCGCTT 4322  
Qy 4709 CAGAGAGGCGAGCTTGTGTCTCAAGAGGAGCATCCCAAGAGCGAAGTGTGAGCGCT 4768  
Db 4323 CAGAGAGGAGGCTTGTGTCTCAAGAGGAGCGTCCCAAGAGCGAAGTGTGATGTACACC 4382  
Qy 4769 GTGAGATGCGCAAGTGTCCCGCACAGACCGTGTCCGAGAGCAGCCACACCCATTCG 4828  
Db 4383 CGGAGAGTGTCCCAAGTGTCCCGCACAGACAGTGTGCGGAGAGCAGCCATCTCC 4442

QY	4829	CCCTTAATGAGCACTGTCTTCGGGGGGGTGATGAGTGGCTGGAACCTGTATTCGACGCAATATCCC	4888
Db	4443	CCCTATATGAGCATTTGTCTCCGGGGGGGTGATGAGTGGTGAACCTGTATTCGATGCAATATCCA	4502
QY	4889	CTGGACCTTTCGACCCCACTCCATATACCCGAGGATTCCTCTGGA---CGAGCCGCTGCC	4945
Db	4503	TTGGACCTTTGACCCCACTCCATATACCCGAGGATTCCTCTGGAATGAGACGACGCGCAGCC	4562
QY	4946	TACTACTGTCCCGACACTTGGCCCCCAACCCCACTTACCCGCACTGTATCCACCTTAC	5005
Db	4563	TACTACTGTCCCGACACTTGGCCCCCAACCCCACTTACCCGCACTGTATCCCACTTAC	4622
QY	5006	CTCATTCGGGGCTTACCCCGACACGGCGGGCGGTGGAGAACCGGACGACATCATTAAGAC	5065
Db	4623	CTCATTCGGGGCTTACCCCGACACGGCGGGCGGTGGAGAACCGGACGACATCATTAAGAC	4682
QY	5066	TACATCACCTTGACGACGATGACACCAACACGGGCAACCGGCACTAGGCGCCAGCGACTGAT	5125
Db	4683	TACATCACCTTGACGACGATGACACCAACACGGGCACTAGGCGCCAGCGACTGAT	4742
QY	5126	ATGCTGAGGGGCTTCTCGCCCGCGAGTCTCTGCTGACCTAACCTAAGCTGCGGGTCCC	5185
Db	4743	ATGCTGAGGGGCTTCTCGCCCGCGAGTCTCTGCTGACCTAACCTAAGCTGCGGGTCCC	4802
QY	5186	CGAGGCATCATGACACTGTCCCAAGTGACACACTGTGCTGTGTGTCGCCCCGACACCA	5245
Db	4803	AGAGGCAATTATCGACCTGTCCCAAGTGACACACTGTGCTGTGTGTCGCCCCGACACCA	4862
QY	5246	GGCACCCCGACGACCGGCATGAGACGGCTTCTACTACCTCCCAACGGCCCGACGCTTC	5305
Db	4863	GGCACCCCTGCGACCGGCATGAGACGGCTTCTACTACCTCCCACTGCGGCCCGACGCTTC	4922
QY	5306	AGCAGCCGCGACAGACAGTCTCCCACTCTCCCCAGAGGTTCACACACTTGAACAAAACA	5365
Db	4923	AGCAGCCGCGACAGAGTACTCACCGGTCTCCCAAGAGGCCCACTCACTTAGTAAACCA	4982
QY	5366	ACCAACACGTCTGTCTCGAGCGGGAGCGAGACCGGATCGAGAGCGGGACCGGGATTCG	5425
Db	4983	ACTCCCACTTTTATCTGAGACGG-----GAAACGGGAACGTGAAGCGG	5024
QY	5426	GAGGGGGGAAAGTCCATCTCAAGTCCACACACGAGTCGTTGAGGACGGACCCATCTGGAGA	5485
Db	5025	GAAAGAGCAAGTCCATCTCAAGTCTCACTACACTAGTGGAGCATGACCCATCTGGAGA	5084
QY	5486	CCTGGTACAGAGCAGAGCAGCGGACGACGCGGACAGCCGCGGGGTGGGGGACGACGC	5545
Db	5085	CCTGGTACGAGAGCAGAGCAGCGGGGCTGGGGGACGACGCG-----	5125
QY	5546	AGCGGCCCGGCTCTCCACTCTCCATGCGCACACGACTGTGCCATTCCTCTCGAACCCAG	5605
Db	5126	-----CCCGGCTCTCCACACCCACGACACTGTGCCATTCCTCTCCGACCCAG	5174
QY	5606	GATGCCCTTCAGAGAGACCCAGTGTGTTTACAACAAGGACATGAAGGATATATCAC	5665
Db	5175	GATGCCCTTCAGAGAGGACCCAGTGTGTTTACAACAAGACATGAAGGAGTGTGATCAC	5234
QY	5666	GCTGTGAGCCAGACAGCCACGGTCTGT-----AGATCCACTCCACTCTCTCA	5716
Db	5235	TCTGTGAGAACCGGACAGCGCCACGGTCTGTAGTGGGGCAAGTTCACCTTCACCTCTTGG	5294
QY	5717	CCCGTTGCCAGCTGCAATTCCTCACTGTGACCCACTGTGCCACTGTGGGCGGCACTCTC	5776
Db	5295	CCTGTGCCCGAGCTGCAATTCCTCACTGTGACCCACTGTGCCACTGTGGGCACTCTT	5354
QY	5777	GATGGGGGTCTAACCTCATATGAGGCCGTCTTGTGTCGCCAAGAGAGCCCCCGGGTC	5836
Db	5355	GAAAGGGGTCTAACCTCATATGAGGCCGTCTTGTGTCGCCAAGAGAACCTCTGGGGTTC	5414
QY	5837	GGCCGGCCAGAGCGGCCCCGAGACACACCGGACATGTCTTCTGTGCAAGCCCCCAGCC	5896
Db	5415	GGCCGGCCAGAGCGGCCCCGGGTGAGCGCTGTGCACTGCTTTTCTTACAAACCCCCGGGC	5474
QY	5897	CGCTTCGGGGTGAAGCCGCGCTCTCTCCCGACCAAGGGCTGGAAGCCCGGCGCTTAGTG	5956

[illegible]



Db	332	CATGTCGGGATATCCACACAGCCTGTGTGGACACAGACGTGGAGGGCCACTGAGCCCGGTATACC	391
Qy	61	GCCCCACAGCCTTTTCTTACCCAGTGGCAGATGCGCCGACCGACACGACGATGCGGCTCTT	120
Db	392	GCCCCACAGCCTTTTCTTACCCAGTGGCAGATGCGCCGACCGACACGACGATGCGGCTCTT	451
Qy	121	GGAGTACGAGACCACTCCCGCGANTATAGCCCTCCACCTGTGTCGGGCTCATTCATCA	180
Db	452	GGAGTACGAGACCACTCCCGGAGTATAGCTCCACCTGTGTCGGGCTCATTCATCA	511
Qy	181	GCCCCAGCGGCGGAGGGCCCTCCCTGTCTGTGAGTTTCAGCCCGGGAAATGAAACGGTCCCA	240
Db	512	GCCCCAGCGGCGGAGGGCCCTCCCTGTCTGTGAGTTTCAGCCCGGGAAATGAAACGGTCCCA	571
Qy	241	GGAAGTCCACCTGCGGGCCAGAGTCCCACTCATACCTGCGCGAGCTGGGAAATGACAGAT	300
Db	572	GGAAGTCCACCTGCGGGCCAGAGTCCCACTCATACCTGCGCGAGCTGGGAAATGACAGAT	631
Qy	301	GGAGTTCAATTGAAAGCAAGCGCCCTTCGGCTAGAGTGTGCTGACCCCTGTCTGCAAC	360
Db	632	GGAGTTCAATTGAAAGCAAGCGCCCTTCGGCTAGAGTGTGCTGACCCCTGTCTGCAAC	691
Qy	361	GTCACCCCTGTCGGGCGACGCGGCGACGCTGCGGGAGCTGAGAGCCCTCACCAAGAACCGAG	420
Db	692	GTCACCCCTGTCGGGCGACGCGGCGACGCTGCGGGAGCTGAGAGCCCTCACCAAGAACCGAG	751
Qy	421	CCTGACGGGCAAGCTGGAACCGGTGTCTCCCCCGACCCCGCCGACACTGACCTTGAGCT	480
Db	752	CCTGACGGGCAAGCTGGAACCGGTGTCTCCCCCGACCCCGCCGACACTGACCTTGAGCT	811
Qy	481	GGAGCTGTGTGCGCGCCACGCGCTGTCCAGAGAGAGCTGATCCAGAACATGACCGGTGTGA	540
Db	812	GGAGCTGTGTGCGCGCCACGCGCTGTCCAGAGAGAGCTGATCCAGAACATGACCGGTGTGA	871
Qy	541	CCGAGAGATTCACCATGTGTAGAGACAGACATCTCTAAGCTGGAAGAAAGAGACGACACACT	600
Db	872	CCGAGAGATTCACCATGTGTAGAGACAGACATCTCTAAGCTGGAAGAAAGAGACGACACACT	931
Qy	601	GGAGGAGAGGCTGTGCAAGCGCGCCCGAGCCTGAGAGAGCCCGGTGTACCGCGCCCATCGA	660
Db	932	GGAGGAGAGGCTGTGCAAGCGCGCCCGAGCCTGAGAGAGCCCGGTGTACCGCGCCCATCGA	991
Qy	661	GTCCGAGCACCGCAGCCTGTGTGCAGATCATCTTACGACGAAACCGGAAGAAAGCTGAAAGC	720
Db	992	GTCCGAGCACCGCAGCCTGTGTGCAGATCATCTTACGACGAAACCGGAAGAAAGCTGAAAGC	1051
Qy	721	TGCACATTCGGGATTCGTGGAAGGCTGTGGGGCCCCAGGTGAGAGCTGCGCTGTTCANCAAC	780
Db	1052	TGCACATTCGGGATTCGTGGAAGGCTGTGGGGCCCCAGGTGAGAGCTGCGCTGTTCANCAAC	1111
Qy	781	CTCCGACACCCGGGCGATTCATGAGAACATCAAAATTAACACAGCGGATGCGGAAGAACT	840
Db	1112	CTCCGACACCCGGGCGATTCATGAGAACATCAAAATTAACACAGCGGATGCGGAAGAACT	1171
Qy	841	AATCTTGTACTTCAAGAGAGGAAATCACGCTCGAAGAACATGGAAGCAAGAAATTCTGCCA	900
Db	1172	AATCTTGTACTTCAAGAGAGGAAATCACGCTCGAAGAACATGGAAGCAAGAAATTCTGCCA	1231
Qy	901	GGCGTATATCCAGAGCTCATGAGAGCCTTGTGAAAAAAAGGTGAGAGCCATTCGAAAAACAACC	960
Db	1232	GGCGTATATCCAGAGCTCATGAGAGCCTTGTGAAAAAAAGGTGAGAGCCATTCGAAAAACAACC	1291
Qy	961	GCGCGCGGGGCGAAGGAGAGCAAGTGTGCGAGATCTACGAAAAAGCAATTCCCTGAGAT	1020
Db	1292	GCGCGCGGGGCGAAGGAGAGCAAGTGTGCGAGATCTACGAAAAAGCAATTCCCTGAGAT	1351
Qy	1021	CCGCAAGCAGCGCGAGCTGCAGAGAGCGCATGCAAGACAGGTGTGGCCAGCGGGGCGACGTGG	1080
Db	1352	CCGCAAGCAGCGCGAGCTGCAGAGAGCGCATGCAAGACAGGTGTGGCCAGCGGGGCGACGTGG	1411
Qy	1081	GCTGTTCATGTGTGCGCGCCCGCAGCGAGACAGAGGTGTTCAGAGATCATGATGCTCTTTC	1140
Db	1412	GCTGTTCATGTGTGCGCGCCCGCAGCGAGACAGAGGTGTTCAGAGATCATGATGCTCTTTC	1471

OY	1141	AGAGCAGAGAACTCGAGAAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATCTGTGA	1200
Db	1472	AGACACAGAGAACTCGAGAAAGCAGATGCGCCAGCTGGCCGTGATCCCGCCCATCTGTGA	1531
OY	1201	CGAGCGTGAACAGAGCGCATCAAGTTTCATCAACATGAACGGGCTTATGCGCGAACCCAT	1260
Db	1532	CGAGCGTGAACAGAGCGCATCAAGTTTCATCAACATGAACGGGCTTATGCGCGAACCCAT	1591
OY	1261	GAAAGTGTACAAAGACCGCCAGGTCATGAACATGTGAGTGAAGAGAGAGAAAGACCTT	1320
Db	1592	GAAAGTGTACAAAGACCGCCAGGTCATGAACATGTGAGTGAAGAGAGAGAAAGACCTT	1651
OY	1321	CCGGGAAAGTTCATGACAGCATCCCAAGAACTTTGGCCCTGATGCCATCATCTCGAGAG	1380
Db	1652	CCGGGAAAGTTCATGACAGCATCCCAAGAACTTTGGCCCTGATGCCATCATCTCGAGAG	1711
OY	1381	GAAACAGTGTGCTGAGTGGCGTCTCTATTAACCTGACTAAGAAAGATGAACCTATATA	1440
Db	1712	GAAACAGTGTGCTGAGTGGCGTCTCTATTAACCTGACTAAGAAAGATGAACCTATATA	1771
OY	1441	GAGCTTGTTGAGACGAGACTATCGGCGCCCGCGCAAGGCCAGACGCAACACAGACGA	1500
Db	1772	GAGCTTGTTGAGACGAGACTATCGGCGCCCGCGCAAGGCCAGACGCAACACAGACGA	1831
OY	1501	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1560
Db	1832	GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1891
OY	1561	AGATGAG	1620
Db	1892	AGATGAG	1951
OY	1621	CGAACAGAGAAACCTCCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1680
Db	1952	CGAACAGAGAAACCTCCCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2011
OY	1681	GAAAGAGGCTGTGGCCTTCAAAGGCCCGCAAACTGCCAACAGCCAGAGAGAGCGCAAGG	1740
Db	2012	GAAAGAGGCTGTGGCCTTCAAAGGCCCGCAAACTGCCAACAGCCAGAGAGAGCGCAAGG	2071
OY	1741	CCGCATCAACCCCGCTCAATATGAGCTTAATGAGGCCCAACAGAGAGAGAGCCATCAACCCCGCAGCA	1800
Db	2072	CCGCATCAACCCCGCTCAATATGAGCTTAATGAGGCCCAACAGAGAGAGAGCCATCAACCCCGCAGCA	2131
OY	1801	GAGCGCCGAGCTGGCCTTCATGAGACTGAATGAGAGTTCTGCTGACAGAGAGAGAGAAAT	1860
Db	2132	GAGCGCCGAGCTGGCCTTCATGAGACTGAATGAGAGTTCTGCTGAGCAGAGAGAGAGAAAT	2191
OY	1861	GGAAACAGCCCAAGAAAGGTCTCTTGAAACACGGCCGCACTGGTGGCCTATGCCCCGAT	1920
Db	2192	GGAAACAGCCCAAGAAAGGTCTCTTGAAACACGGCCGCACTGGTGGCCTATGCCCCGAT	2251
OY	1921	GGTGGGCTCCAAAGCTGTGTGCAAGTGAAGAACTTCACTTCAACTCAAGAGAGAGCA	1980
Db	2252	GGTGGGCTCCAAAGCTGTGTGCAAGTGAAGAACTTCACTTCAACTCAAGAGAGAGCA	2311
OY	1981	GAACTCGATGAGATCTTTCAGCAGCAACAGCTGAAGATGAGAGAGAGAGAGAAACGCGCG	2040
Db	2312	GAACTCGATGAGATCTTTCAGCAGCAACAGCTGAAGATGAGAGAGAGAGAGAGAAACGCGCG	2371
OY	2041	GAGGAAAGAAAGAAAGCGCCGCGCGCGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2100
Db	2372	GAGGAAAGAAAGAAAGCGCGCGCGCGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG	2431
OY	2101	GGAAGATGAGAGATGAGAGCGTGGGCGTGTGAGCGGAAATAGAGAGAGAGATGTGTGAGGA	2160
Db	2432	GGAAGATGAGAGATGAGAGCGTGGGCGTGTGAGCGGAAATAGAGAGAGAGATGTGTGAGGA	2491
OY	2161	GGCTGAAGCCTTACATGCTCTTGAGGAATGAGGTGCCAGAGAGGGAATGCAATGCTCCAGC	2222
Db	2492	GGCTGAAGCCTTACATGCTCTTGAGGAATGAGGTGCCAGAGAGGGAATGCAATGCTCCAGC	2551

2221 CACTGTCAACACAGCTCAGACACCGAGAGATCCCTCTCTCACTGAGGCCGCA 2280  
2552 CACTGTCAACACAGCTCAGACACCGAGAGATCCCTCTCTCACTGAGGCCGCA 2611  
2281 GGAACACAGGCGAGATGGGCGCAAGCCCGCAGCCACTTGGGGCGGAGGCGACCC 2340  
2612 GGAACACAGGCGAGATGGGCGCAAGCCCGCAGCCACTTGGGGCGGAGGCGACCC 2671  
2341 AGGCCCCACCCACCCACCGAGGACATCCGGGGCCCCATTGAGCCACCCGGGCT 2400  
2672 AGAGCCACCCACCCACCGAGGACATCCGGGGCCCCATTGAGCCACCCGGGCT 2731  
2401 TGAAGCACCGAGGCCCTTACGCGCCCGCAGCAGCAGCCCATCGCCCTTGCACCTCC 2460  
2732 TGAAGCACCGAGGCCCTTACGCGCCCGCAGCAGCAGCCCATCGCCCTTGCACCTCC 2791  
2461 TGTGTCTCCCAAG 2520  
2792 TGTGTCTCCCAAG 2851  
2521 GGAAGGAG 2580  
2852 GGAAGGAG 2911  
2581 GCCCGCTCAAG 2640  
2912 GCCCGCTCAAG 2971  
2641 GGAAGGCGCTGAG 2700  
2972 GGAAGGCGCTGAG 3031  
2701 GGAAGGAG 2760  
3032 GGAAGGAG 3091  
2761 CTGCACTGACAG 2820  
3092 CTGCACTGACAG 3151  
2821 AAGGCCAGAGCTTCTCAACCCCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880  
3152 AAGGCCAGAGCTTCTCAACCCCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3211  
2881 ACTGAGACTGAG 2940  
3212 ACTGAGACTGAG 3271  
2941 AGTCATGAG 3000  
3272 AGTCATGAG 3331  
3001 GCCACCGCAAAACCTGACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060  
3332 GCCACCGCAAAACCTGACGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3391  
3061 GGGCAAG 3100  
3392 GGGCAAG 3451  
3101 ----GCGTTGCGAGCGAG 3456  
3452 CCCAGGCTTTCAGAGCGAG 3511  
3157 CCGGCGCTTTCAGAGCGAG 3216  
3512 CCGGCGCTTTCAGAGCGAG 3571  
3217 CTGAGGCTTTCAGAGCGAG 3276  
3572 CTGAGGCTTTCAGAGCGAG 3631  
3277 CCGGCGGCTCTGCGGAG 3336

3632 CCAGGCGGCTCTGCGGAG 3691  
3337 CAAGCAGCCAGAGCTTCAG 3396  
3692 CAAGCAGCCAGAGCTTCAG 3751  
3397 GCTCCAGCTCCGTACTCAG 3456  
3752 GCTCCAGCTCCGTACTCAG 3811  
3457 GCCCTGCTCCAG 3516  
3812 GCCCTGCTCCAG 3871  
3517 GTCCCGAG 3576  
3872 GTCCCGAG 3931  
3577 GTCCCGAG 3636  
3932 GTCCCGAG 3991  
3637 TTCAG 3696  
3992 TTCAG 4051  
3697 CAGCGAG 3756  
4052 CAGCGAG 4111  
3757 GAGTCGCTTGAAG 3816  
4112 GAGTCGCTTGAAG 4171  
3817 CAAG 3876  
4172 CAAG 4231  
3877 GAGCGAG 3936  
4232 GAGCGAG 4291  
3937 CATGATGAG 3996  
4292 CATGATGAG 4351  
3997 CCGTGCATTCGCGCGAG 4056  
4352 CCGTGCATTCGCGCGAG 4411  
4057 CCGGTCATTCAG 4116  
4412 CCGGTCATTCAG 4471  
4117 TCGGAG 4176  
4472 TCGGAG 4531  
4177 CTTGACAG 4236  
4532 CTTGACAG 4591  
4237 GGGGCTGAG 4296  
4592 GGGGCTGAG 4651  
4297 GCTGAG 4331  
4652 GCTGAG 4686

RESULT 10



COJ22208 LOCUS COJ22208 6339 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 8142 from Patent WO02068579.  
ACCESSION COJ22208  
VERSION COJ22208.1 GI:42283065  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE  
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
Kites, such as nucleic acid arrays, comprising a majority of  
humanecons or transcripts, for detecting expression and other uses  
hereof  
JOURNAL Patent: WO 02068579-A 8142 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source  
1..6339  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 46.3%; Score 3963.8; DB 6; Length 6339;  
Best Local Similarity 75.6%; Pred. No. 0;  
Matches 5911; Conservative 0; Mismatches 167; Indels 1737; Gaps 11;  
2 ATGTGGGGCTCCACACAGTTGTGGACAGAGTGAAGGGCCACTGAGCCCGCTACCCG 61  
1 ATGTGGGATTCACACAGCTGTGGACAGAGTGAAGGGCCACTGAGCCCGCTACCCG 60  
62 CCCACAGGCTTCTCTACCCAGTGCAGATGCGCCGACGACACAGGACGTCGGGCTCTG 121  
61 CCCACAGGCTTCTCTACCCAGTGCAGATGCGCCGACGACACAGGACGTCGGGCTCTG 120  
122 GAGTACAGACACACTCCCGACACTATGCTCCACCTGTGCGCGGGCTCCATCATTCAG 181  
121 GAGTACAGACACACTCCCGACACTATGCTCCACCTGTGCGCGGGCTCCATCATTCAG 180  
132 CCCGACGGCGGAGGCGCTCTGCTGTGTGAGTTCCAGCCGGGAATGAAGGCTCCAG 241  
131 CCCGACGGCGGAGGCGCTCTGCTGTGTGAGTTCCAGCCGGGAATGAAGGCTCCAG 240  
242 GAGCTCCACTGCGGCGAGAGTCCCACTGATACCTGCGGAGTGGGGAATCAAGATG 301  
241 GAGCTCCACTGCGGCGAGAGTCCCACTGATACCTGCGGAGTGGGGAATCAAGATG 300  
302 GAGTTCATTGAAGAAGGAGCGCTCGGCTAGAGCTGCTGACCCCTGCTGCGACCG 361  
301 GAGTTCATTGAAGAAGGAGCGCTCGGCTAGAGCTGCTGACCCCTGCTGCGACCG 360  
362 TCAACCTCTGCTGCGACCGGCGACGCTGCGGATCTGAAGACCTCAACAGACCGTAGC 421  
361 TCAACCTCTGCTGCGACCGGCGACGCTGCGGATCTGAAGACCTCAACAGACCGTAGC 420  
422 CTGAGGGGCAAGCTGGAACCGGTGTCTCCCGCAAGCCCCCGCACTGAACCTTAGGCTG 481  
421 CTGAGGGGCAAGCTGGAACCGGTGTCTCCCGCAAGCCCCCGCACTGAACCTTAGGCTG 480  
482 GAGCTGTGCGCCGACGCTGTCCAAAGAGAGTGAATCCAGAACATGAGCCGCTGAC 541  
481 GAGCTGTGCGCCGACGCTGTCCAAAGAGAGTGAATCCAGAACATGAGCCGCTGAC 540  
542 CGAGAGATCAACATGTGTAGAGACAGATCTCTAAGCTGAAGAAGAGACAGCAAGCTG 601  
541 CGAGAGATCAACATGTGTAGAGACAGATCTCTAAGCTGAAGAAGAGACAGCAAGCTG 600  
602 GAGGAGGAGGCTGCGCAAGCGCGCGAGCCTTGAGAAAGCCGCTGACCCGCGCCATCGAG 661  
601 GAGGAGGAGGCTGCGCAAGCGCGCGAGCCTTGAGAAAGCCGCTGACCCGCGCCATCGAG 660

QY 662 TCGAAGCACCGGAGCTGTGTGACATCTACGACGAGAACCGGAGAGAGGCTGAGCT 721  
DB 661 TCGAAGCACCGGAGCTGTGTGACATCTACGACGAGAACCGGAGAGAGGCTGAGCT 720  
QY 722 GCACATCGGATCTTGAAGGCTGTGGGCGCCGAGTGAAGCTGCGCTGTACCAACGAGCC 781  
DB 721 GCACATCGGATCTTGAAGGCTGTGGGCGCCGAGTGAAGCTGCGCTGTACCAACGAGCC 780  
QY 782 TCCGACACCGGCACTATCATGAGAACATCAAAATAAACAGGCGATCGGAGAGCTA 841  
DB 781 TCCGACACCGGCACTATCATGAGAACATCAAAATAAACAGGCGATCGGAGAGAGCTA 840  
QY 842 ATCTGTACTTAAAGAGAGAAATGACGCTGTGGAAACAATGGAACCAAGATTTCCGAG 901  
DB 841 ATCTGTACTTAAAGAGAGAAATGACGCTGTGGAAACAATGGAACCAAGATTTCCGAG 900  
QY 902 CGCTATGACCAAGCTATGAGAGGCTTGGAAAAAAGTGAAGCGATGGAACCAACCG 961  
DB 901 CGCTATGACCAAGCTATGAGAGGCTTGGAGAGAGAGTGAAGCGATGGAACCAACCG 960  
QY 962 CGCCGCGGCGCAAGAGAGCAAGGTGCGGAGTACTGAAAAAGCAGTTCCTGAGATC 1021  
DB 961 CGCGCGCGGCGCAAGAGAGCAAGGTGCGGAGTACTGAAAAAGCAGTTCCTGAGATC 1020  
QY 1022 CGCAAGCAGCGGAGCTGAGAGAGCGCATTCAGACAGAGGTGGCTCAGCGGCGAGTGG 1081  
DB 1021 CGCAAGCAGCGGAGCTGAGAGAGCGCATTCAGAGAGCGCATTCAGAGAGCGGCGAGTGG 1077  
QY 1082 CTGTCAATGTGGGCGCGCGAGAGAGCAAGGTGAGAGATGATGATGAGTGGCTCTCA 1141  
DB 1078 CTGTCAATGTGGGCGCGCGAGAGAGCAAGGTGAGAGATGATGATGAGTGGCTCTCA 1137  
QY 1142 GAGCAGAGAACTTGAAGAGAGATGCGGACAGTGGCGGTGATCCCGCCATGCTGTAC 1201  
DB 1138 GAGCAGAGAACTTGAAGAGAGATGCGGACAGTGGCGGTGATCCCGCCATGCTGTAC 1197  
QY 1202 GAGCTGACAGAGCGCATCAAGTTCAATCAACATGAACGAGCTTATGAGCGACCCATG 1261  
DB 1198 GAGCTGACAGAGCGCATCAAGTTCAATCAACATGAACGAGCTTATGAGCGACCCATG 1257  
QY 1262 AAGGTGCAAAAGACCGGACAGTGCATGAAACATGAGAGAGAGAGAGACCTTC 1321  
DB 1258 AAGGTGCAAAAGACCGGACAGTGCATGAAACATGAGAGAGAGAGAGACCTTC 1317  
QY 1322 CGGAGAACTTCAATCAACATCCAAAGACTTGTGCTGATGCTCATCTTCTGAGAGG 1381  
DB 1318 CGGAGAACTTCAATCAACATCCAAAGACTTGTGCTGATGCTCATCTTCTGAGAGG 1377  
QY 1382 AAGACAGTGGCTGAGTGGCTCTTATTACTGACTAAGAGAAATGAGAACTATAAG 1441  
DB 1378 AA----- 1379  
QY 1442 AGCTGTGAGACGAGAGCTATCGGCGCGGCAAGACCGACAGCAACAGACGAG 1501  
DB 1380 -----GATGAGTGTGCTGCGCGCCCATACCTTC 1409  
QY 1502 CAGCAGACAGCAGACAGACAGACAGACGACCCCGACAGACGACGAGAGAGAAA 1561  
DB 1410 GGTCTTCACTTCGTGGGCACTGCTGAGCTCAG----- 1445  
QY 1562 GATGAGAGAGAGAGAGAGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1621  
DB 1446 ----- 1445  
QY 1622 GACAAAGAAAGACTTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1681  
DB 1446 -----GAG 1485  
QY 1682 AAGGAGGCTGTGGCTTCAAAAGCGCGCAAACTGCAACAGCAGAGAGAGAGAGAGAG 1741  
DB 1486 AAGGAGGCTGTGGCTTCAAAAGCGCGCAAACTGCAACAGCAGAGAGAGAGAGAGAG 1545  
QY 1742 CGCATCAACCGCTCAATGCTTAATGAGGCAACAGCAGAGAGAGAGAGAGAGAGAG 1801

1546 CGCATCACCGCTCAATGGCTAATGAGCCACAGCGAGGCGCATCACCCCCAGAG 1605  
1802 AGGCGGAGGTGGCTCTCCATGAGAGCTGAAATGAGATTCTGCTGAGAAAGAAATG 1861  
1606 AGGCGGAGGTGGCTCTCCATGAGAGCTGAAATGAGATTCTGCTGAGAAAGAAATG 1665  
1862 GAACAAGCCAAAGAAAGTCTCTCTGAAACAGCGCGCAACTGTCTGCGCATCGCCGATG 1921  
1666 GAACAAGCCAAAGAAAGTCTCTCTGAAACAGCGCGCAACTGTCTGCGCATCGCCGATG 1725  
1922 GTGGGCTCCAAAGCTGTGTGCGAGTGTAAAGCTTCTACTTCACTAACAAGAGGCGAG 1981  
1726 GTGGGCTCCAAAGCTGTGTGCGAGTGTAAAGCTTCTACTTCACTAACAAGAGGCGAG 1785  
1982 AACCTGATGATGATCTTTGACAGACACAACTGAAATGAGAAAGAAAGAAAGCGCGG 2041  
1786 AACCTGATGATGATCTTTGACAGACACAACTGAAATGAGAAAGAAAGAAAGCGCGG 1845  
2042 AGGAAGAAAGAAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 2101  
1846 AGGAAGAAAGAAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 1905  
2102 GAGGATGAGAGATGAGAGCGCTCGGCGGTGAGCGGAAATGAGAGAGATGTGAGAGAG 2161  
1906 GAGGATGAGAGATGAGAGCGCTCGGCGGTGAGCGGAAATGAGAGAGATGTGAGAGAG 1965  
2162 GCTGAAGCTTAACTGCTCTTGGGAATGAGGTGCCAGAGGGGAAATGCAATGGCCAGCC 2221  
1966 GCTGAAGCTTAACTGCTCTTGGGAATGAGGTGCCAGAGGGGAAATGCAATGGCCAGCC 1993  
2222 ACTGTAAACAAGCTCAGACACGAGAGATCCCTCTCTCTCAACTGAGGCGGCGAAG 2281  
1994 ----- 1993  
2282 GACACAGGCGAGATGGGCCCAAGCCCCCAAGCACTCTGGCGCGAGCGGCGCAACCCCA 2341  
1994 ----- 1993  
2342 GGGCCACCAACCCCAACGAGAGACATCCCGGCCCCCATTTGAGCCACCCCGGCTCT 2401  
1994 ----- 1993  
2402 GAAGCAACGAGAGCCCTTACGCCCCCAACAGCAACCCCATCGCCTCTGCACTCTCT 2461  
1994 ----- 1993  
2462 GTGGTCCCAAGAGAGAGAAAGAGAGAGACCGCAGACGCGCCCACTGAGAGAGGG 2521  
1994 ----- 1993  
2522 GAGGAGCAAGAGCCCCCGCGCTGAGAGAGCTGCACTGAGCAACAGGAAAGCCGAGAGAG 2581  
1994 ----- 1993  
2582 CCCGTCAAGAGAGTGCACGAGAGAAAGCCGAGAGGGGCGCGCAAGGCGAAGACGG 2641  
1994 ----- 1993  
2642 GAGGCGCTGAGGCGCACGCGCGAGGGGGCGCTCAAGGCAAGAAAGAGAGGGCGGAGC 2701  
1994 ----- 1993  
2702 GCGAGGCGCACCTGCGCAAGAGCTGGGCGCGCCCCCAGAGCAAGCGACTCACTGCTTACC 2761  
1994 ----- 1993  
2762 TGCAATGACAGAGAGTGAATGAGCGAGAGGGGCGAGCAAGAACGGGCTGTCTCCCA 2821  
2017 TGCAATGACAGAGAGTGAATGAGCGAGGGCGGCGAGCAAGAACGGGCTGTCTCCCA 2076  
2822 AGGCGGAGCTCTCAACCCGACTGGCGACCCCGGCGCAATGCTCAACCCGAGAGCCA 2881

2077 AGGCGGAGCTCTCTCAACCCGACTGGCGACCCCGGCGCAATGCTCAACCCGAGAGCCA 2136  
2882 CTGGAACCTGAAGAGCTGAAGAGCAGCGAGGCTGCGCATTCGCCCATCAAGTACCAAA 2941  
2137 CTGGAACCTGAAGAGCTGAAGAGCAGCGAGGCTGCGCATTCGCCCATCA 2184  
2942 GTCCATGAGGCCCGCGGAGAGCAGAGCTCCCAACCAAGCACTGCCAGCCCAACCG 3001  
2185 ----- 2184  
3002 CCACCGCAAAACCTGCAACCCGAGAGAGACGCCCTTCAGAGCTTGACAGACGCCCGG 3061  
2185 ----- 2184  
3062 GGCAGAGCAGAGAGCGCGGACACCCCGCGCAAGAGGCTTTCGACGCGAGGCCAG 3121  
2185 ----- 2191  
3122 AAGCTGCTGGAGACCCCTTGTGAACTTCCGCGCTGCTTCCCGTCCCGCCCGT 3181  
2192 ACCGCTCTGGGCGCCCAATGAGACAGACTGG----- 2225  
3182 GAGGTATCAAGGCTTCCCGCATGCCCGGACCCCTCAAGCTTCTCTAAGCTCACT 3241  
2226 ----- 2232  
3242 GGTCAACCACTGGCGCTTGGGCTTCAATGACATGCGCGCGCGCTGCGCGCGCGCAAC 3301  
2233 GGTCAACCACTGGCGCTTGGGCTTCAATGACATGCGCGCGCGCTGCGCGCGCGCAAC 2292  
3302 ACCATCTCAACCGCGCTCCCTCATCTCTCTGCAAGACCCAGCGTCTGAGAGG 3361  
2293 ACCATCTCAACCGCGCTCCCTCATCTCTCTGCAAGACCCAGCGTCTGAGAGG 2352  
3362 CAATAGTGCATCTTCCAGAGAAATGTGCTCAAGCTCAAGTCTCCCTGATCTAGAGCAT 3421  
2353 CAATAGTGCATCTTCCCA----- 2373  
3422 GCCAAGGCCCGGTGGGCGCTGTCAACATGGGGCTGCCCTGCGCATGAGACCCCAAAA 3481  
2374 ----- 2373  
3482 CTGGCAACCTTCAAGCGAGTGAAGCAGAGCACTGTCCCAAGGGCGAGGCTGGGCA 3541  
2374 ----- 2373  
3542 CCGAGAGCCTGGGGTGCACACAGGCCAGAGGCGTCTGCTGAGAGGAGCACTCTG 3601  
2374 ----- 2373  
3602 GGCTCAAGTTCGGGCGGAAAGCATCAACMAAGGATTCACAGCACAGGAGTCCCTGGAG 3661  
2374 ----- 2373  
3662 AGGCGCATCAATACCGGCGCTCATCAACCAAGGACGCGCAGCTGACTGTACAAG 3721  
2374 ----- 2400  
3722 GGCACCATCAACAGATCATTCGCGAGAGACAGCCGAGTGTGCTTGAACCGCGCGGAG 3781  
2401 GGCACCATCAACAGATCATTCGCGAGAGAGAGCCGAGTGTGCTTGAACCGCGCGGAG 2460  
3782 GACAGCTGCGCAAGGCGCACTCATCTACAGAGGCAAGAGGCGCAAGTGTCTTCTAT 3841  
2461 GACAGCTGCGCAAGGCGCACTCATCTACAGAGGCAAGAGGCGCAAGTGTCTTCTAT 2520  
3842 GAGGCTGCGCATGTCTGACCCAGTCTCAAGAGAGAGCGGCAAGAGCAAGCTCAAGACC 3901  
2521 GAGGCTGCGCATGTCTGACCCAGTCTCAAGAGAGAGCGGCAAGAGCAAGCTCAAGACC 2580  
3902 CCCCATGAGACGCGCGCGCCCAAGCGCACCTATGACATGAGAGGGCGCGTGGCAAG 3961  
2581 CCCCATGAGACGCGCGCGCCCAAGCGCACCTATGACATGAGAGGGCGCGTGGCAAG 2640

QY 3962 GGCATCTCTCAGCCAGCATCGAAGTCTCATGAGCCGTGCATCCCGCCGAGCGACAC 4021  
DB 2641 GGCATCTCTCAGCCAGCATCGAAGT----- 2667  
QY 4022 AGCCCCCACCTCAMAAGACGACACATCCGGGCTCCATCACACAGAGATCCCT 4081  
DB 2668 ----- 2667  
QY 4082 CGGTCTTACGTGAGGACACAGAGGACTACGTGCGTCGGAGGCCAAGCTCTAAAGCGG 4141  
DB 2668 ----- 2667  
QY 4142 GAGGACAGCTTCGCCCCCAACGCGCTTCACGAGACCTTGACGAGGCTTACAAAGACGACG 4201  
DB 2668 GATTACAGAGGAGAGACTTTCATCTCGGTGCCCCCTGTGGGCGGCTTACAAAGACGACG 2727  
QY 4202 GCGCTGGGCCCCCTGAGAGCTGAAGCCGAGCTTGAAGGCTTGTGGCCACGGTGAAGAG 4261  
DB 2728 GCGCTGGGCCCCCTGAGAGCTGAAGCCGAGCTTGAAGGCTTGTGGCCACGGTGAAGAG 2787  
QY 4262 GCGGAGCGCTCATCATGATGATCCCGGAGGAGAGCTGCGGACACGCGCGAGCTGCC 4321  
DB 2788 GCGGAGCGCTCATCATGATGATCCCGGAGGAGAGCTGCGGACACGCGCGAGCTGCC 2847  
QY 4322 CTGAGCCCGCGGCGCTCAAGAGAGGCTTCATCACGACAGGACACCCGCTCAAGTACAC 4381  
DB 2848 CTGAGCCCGCGGCGCTCAAGAGAGGCTTCATCACGACAGGACACCCGCTCAAGTACAC 2907  
QY 4382 ACCGCGCGCTCCACCATCTGAGCTTCAAAAAAGACAGACTACGCTCCCTCATCGGACGCC 4441  
DB 2908 ACCGCGCGCTCCACCATCTGAGCTTCAAAAAAGACAGACTACGCTCCCTCATCGGACGCC 2967  
QY 4442 GCGCGGAGCTTCCACACCGGTGACACCGCTGAGTGTGATGAGCGACGCGCGGACCTGGA 4501  
DB 2968 GCGCGGAGCTTCCACACCGGTGACACCGCTGAGTGTGATGAGCGACGCGCGGACCTGGA 3027  
QY 4502 CGTGCCTGCTACGAGAGAGAGCTTGAAGACCGGACGAGGACCGGACGCTTCGAGGAGC 4561  
DB 3028 CGTGCCTGCTACGAGAGAGAGCTTGAAGACCGGACGAGGACCGGACGCTTCGAGGAGC 3087  
QY 4562 TCCATTGCGCGGCGGCGCGCTCATTTGCTGAGCTGAGTGAAGCCGCGGACAGCCCC 4621  
DB 3088 TCCATTGCGCGGCGGCGCGCTCATTTGCTGAGCTGAGTGAAGCCGCGGACAGCCCC 3147  
QY 4622 CTGACTTATGAGAGACACGAGGAGACCTTTGCGGACCTCCACAGAGTTGCGCGGTG 4681  
DB 3148 CTGACTTATGAGAGACACGAGGAGACCTTTGCGGACCTCCACAGAGTTGCGCGGTG 3207  
QY 4682 ACCATGCGGAGACCCACGCGCGGCTTGACGAGAGGAGCGCTTTGCTACAGCAAGGATCC 4741  
DB 3208 ACCATGCGGAGACCCACGCGCGGCTTGACGAGAGGAGCGCTTTGCTACAGCAAGGATCC 3267  
QY 4742 CAGGACCGAAGAGCTGACGTGACGCGCTGTGAGATGCGCAAGTCCCGGACAGCACCGTG 4801  
DB 3268 CAGGACCGAAGAGCTGACGTGACGCGCTGTGAGATGCGCAAGTCCCGGACAGCACCGTG 3327  
QY 4802 CCGGAGACACACCCACACCTTTCGACCTTGAAGACCTGCTTCGAGGCGGTGAGTGAGC 4861  
DB 3328 CCGGAGACACACCCACACCTTTCGACCTTGAAGACCTGCTTCGAGGCGGTGAGTGAGC 3387  
QY 4862 GTTGAACCTGTATCGAGACCATCCCGCTGAGCGCTTTCGACCCGACCTTACACCGCGGAC 4921  
DB 3388 GTTGAACCTGTATCGAGACCATCCCGCTGAGCGCTTTCGACCCGACCTTACACCGCGGAC 3447  
QY 4922 ATCCCTCTTGAAGACGAGCGCTGCTTATCTGACCTGCGGACACCTTGGCCCCAACCCACCC 4981  
DB 3448 ATCCCTCTTGAAGACGAGCGCTGCTTATCTGACCTGCGGACACCTTGGCCCCAACCCACCC 3507  
QY 4982 TACCGGACCTTGAAGACGAGCGCTTATCTGACCTGCGGAGTACCGGACACGAGGCGGCTGAG 5041  
DB 3508 TACCGGACCTTGAAGACGAGCGCTTATCTGACCTGCGGAGTACCGGACACGAGGCGGCTGAG 3567

QY 5042 AACCGGACGACCATCATGATGATGATCATCATCTTCGACGACGATGACACCAACACGCGCC 5101  
DB 3568 AACCGGACGACCATCATGATGATGATGATCATCATCTTCGACGACGATGACACCAACGCGCC 3627  
QY 5102 ACCGCGAGGCGCCGAGGAGCTGATATGCTGAGGAGGCGCTTGGCCCCGAGGCTTCGCGTG 5161  
DB 3628 ACCGCGAGGCGCCGAGGAGCTGATATGCTGAGGAGGCGCTTGGCCCCGAGGCTTCGCGTG 3687  
QY 5162 GCACTCAACTAGCGCTGCGGAGTCCCGAGGAGCATCATGACCTGTCCCAAGTGACACCTTG 5221  
DB 3688 GCACTCAACTAGCGCTGCGGAGTCCCGAGGAGCATCATGACCTGTCCCAAGTGACACCTTG 3747  
QY 5222 CCTGTGCTGTGCCCCGACACACGAGCACCCGAGCACCGGCATGAGACCGCTTGTCTAC 5281  
DB 3748 CCTGTGCTGTGCCCCGACACACGAGCACCCGAGCACCGGCATGAGACCGCTTGTCTAC 3807  
QY 5282 CTCCCAACCGGCCCCCAACCCCTTCAAGAGCGGACACAGAGTCCCACTGCCCAAGA 5341  
DB 3808 CTCCCAACCGGCCCCCAACCCCTTCAAGAGCGGACACAGAGTCCCACTGCCCAAGA 3867  
QY 5342 GATCCACACACTTTCACAAAACCAACCAACAGTCTCTGTCCGAGCGGAGGAGACCGG 5401  
DB 3868 GATCCACACACTTTCACAAAACCAACCAACAGTCTCTGTCCGAGCGGAGGAGACCGG 3927  
QY 5402 GATTCGAGCGGAGACCGGAGATCGGAGAGCGGAAAAGTCCATCTCTACCTCCACACGACG 5461  
DB 3928 GATTCGAGAGCGGAGACCGGAGATCGGAGAGCGGAAAAGTCCATCTCTACCTCCACACGACG 3987  
QY 5462 GTGAGGACGAGACCCATCTGAGAGACTGTGATACAGAGACGAGGAGCGGACCGGACG 5521  
DB 3988 GTGAGGACGAGACCCATCTGAGAGACTGTGATGAGGAGATACAGAGCCCAACCCCGCTCCG 4047  
QY 5522 AGCGCGGAGGAGTGGGAGGACGAGGACGCGCGCGCTCCACTCCATGACCAACGAGAC 5581  
DB 4048 GGACTCTTGTGAGGCGGAGAGAGGCGCTCCCGTGTGATGACATGACATGACACGAGAC 4107  
QY 5582 TCGCCCATCTCCCTCGGACCCAGAGATGCTTCCAGCAGAGACCCAGATGCTTTACAC 5641  
DB 4108 TCGCCCATCTCCCTCGGACCCAGAGATGCTTCCAGCAGAGACCCAGATGCTTTACAC 4167  
QY 5642 ACAGGCAAGAGGAGTATCATACCGCTGTGAGAGCCCAACAGCCACGAGTCCGAGGTCC 5701  
DB 4168 ACAGGCAAGAGGAGTATCATACCGCTGTGAGAGCCCAACAGCCACGAGTCCGAGGTCC 4227  
QY 5702 ACTTCACCTCTCAACCCGTTTCGACCACTGACATTTCCACCTTGACACCACTGACCA 5761  
DB 4228 ACTTCACCTCTCAACCCGTTTCGACCACTGACATTTCCACCTTGACACCACTGACCA 4287  
QY 5762 CTGAGCGGACCCCTGATGAGGAGTCTAACCTTACCTCATGAGAGCCGTTTGTGCGCCAG 5821  
DB 4288 CTGAGCGGACCCCTGATGAGGAGTCTAACCTTACCTCATGAGAGCCGTTTGTGCGCCAG 4347  
QY 5822 GAGGCGCGCGGAGTTCGCGGCAAGAGGCGCCGAGGACACCGGACATGAGCTTCCG 5881  
DB 4348 GAGGCGCGCGGAGTTCGCGGCAAGAGGCGCCGAGGACACCGGACATGAGCTTCCG 4407  
QY 5882 GCGAAGCCCGGACCGGCTCGGAGCTGAGGCGCGGCTCTCCCGGACGAGGAGCTCGAG 5941  
DB 4408 GCGAAGCCCGGACCGGCTCGGAGCTGAGGCGCGGCTCTCCCGGACGAGGAGCTCGAG 4467  
QY 5942 CCGCGGCGCTTATGATGCTCTGTCTTGGCCACGCGACCATGCGCGGACACCCCTGCGAAG 4527  
DB 4468 CCGCGGCGCTTATGATGCTCTGTCTTGGCCACGCGACCATGCGCGGACACCCCTGCGAAG 4587  
QY 6002 AACCTTCGACCTTACACAGGACCGCGGACCGCGGCGGACCTGAGGCGGCTCGAGC 6061  
DB 4588 AACCTTCGACCTTACACAGGACCGCGGACCGCGGCGGACCTGAGGCGGCTCGAGC 4587  
QY 6062 CCGGACCGGAGAAAGACTCAAAAGTAAACCTTTTCATTCAGAGAACTGAACTCGGTTCT 6121  
DB 4588 CCGGACCGGAGAAAGACTCAAAAGTAAACCTTTTCATTCAGAGAACTGAACTCGGTTCT 4647  
QY 6122 CTG----- 6124

Db	4648	CTGGGTAAGACCA	CCCTTGA	CAGCGGCAC	CTTCA	TAGACGGATTAT	CATGCTCAATT	4707
Qy	6125	-----	-----	-----	-----	-----	-----	6127
Db	4708	GCTCAGATAAAG	GGGCGCGA	GAAGAGTG	CGCTG	GCACAA	CGGCTCCCTCG	CGATGT
Qy	6128	TACCA	CGGCACAG	CTTAC	AGCCCCG	AAAGGGT	GGAGCCCG	CTGTAGCTTAC
Db	4768	TACCA	CGGCACAG	CTTAC	AGCCCCG	AAAGGGT	GGAGCCCG	CTGTAGCTTAC
Qy	6188	AGT	CTGAC	CCCA	CGACAG	AGGGCT	CCCCA	AGCACTGGA
Db	4828	AGT	CTGAC	CCCA	CGACAG	AGGGCT	CCCCA	AGCACTGGA
Qy	6248	GAG	GGGAG	CTG	CGGCG	CCCA	AGCAG	CGCCCGTGA
Db	4888	GAG	GGGAG	CTG	CGGCG	CCCA	AGCAG	CGCCCGTGA
Qy	6308	CT	CCCA	CA	CTG	CGGCG	CCCTG	CTGAC
Db	4948	CT	CCCA	CA	CTG	CGGCG	CCCTG	CTGAC
Qy	6368	G	CCCA	CGG	GGT	CA	AGG	CTC
Db	5008	G	CCCA	CGG	GGT	CA	AGG	CTC
Qy	6422	-----	-----	-----	-----	-----	-----	6421
Db	5068	AC	GGGTT	CT	G	CTG	CTG	CTG
Qy	6422	-----	-----	-----	-----	-----	-----	6421
Db	5128	AT	GGGAA	CC	CC	AG	CC	CTT
Qy	6422	-----	-----	-----	-----	-----	-----	6421
Db	5188	CT	GGG	GGG	AG	CT	CA	AGG
Qy	6461	C	AG	CT	CA	AGG	CT	CA
Db	5248	C	AG	CT	CA	AGG	CT	CA
Qy	6521	CT	GA	AC	CT	CG	CG	CG
Db	5308	CT	GA	AC	CT	CG	CG	CG
Qy	6581	G	CC	CT	G	CT	G	CT
Db	5368	G	CC	CT	G	CT	G	CT
Qy	6641	G	CT	T	G	G	T	G
Db	5428	G	CT	T	G	G	T	G
Qy	6701	C	CA	G	G	CA	CT	C
Db	5488	C	CA	G	G	CA	CT	C
Qy	6761	C	CA	G	G	CA	CT	C
Db	5548	C	CA	G	G	CA	CT	C
Qy	6821	A	AG	CT	G	A	AG	CT
Db	5605	A	AG	CT	G	A	AG	CT
Qy	6881	C	T	GA	A	C	C	A
Db	5665	C	T	GA	A	C	C	A
Qy	6941	T	T	C	A	T	A	T

Db	5725	TTCAATATATGCCCGGCATCAACCGAACAAGGCTTATGACTATAGAAAGCAAGCGGTGAC	5788
Qy	7001	GAACATGCGACGACCAACATATGGGGCTGAGGGCCATATTAGAAAGCACTCATGGTTAA	7060
Db	5785	GAACATGCGACGACCAACATATGGGGCTGAGGGCCATATTAGAAAGGCACTCATGGTTAA	5844
Qy	7061	TATGACCAAGTGGGAAGATCCCGCGCTGACGGGCAATGCTTTTAAACCTCTGAATGCC	7120
Db	5845	TATGACCAAGTGGGAAGATCCCGCGCTGACGGGCAATGCTTTTAAACCTCTGAATGCC	5904
Qy	7121	AGTGCCAGGCTGCGCGCTGCTATGCCCCATTAACGCTGCTGAACGACGAGTGACCAACA	7180
Db	5905	AGTGCCAGGCTGCGCGCTGCTATGCCCCATTAACGCTGCTGAACGACGAGTGACCAACA	5966
Qy	7181	CTCACCTTGCCAGATGGTGGCGGCGGGGAAGGCCAAGTCTCTGCGAAGCCACGACCGGAAA	7240
Db	5965	CTCACCTTGCCAGATGGTGGCGGCGGGGAAGGCCAAGTCTCTGCGAAGCCACGACCGGAAA	6024
Qy	7241	GCCAAAGTCCCCGGGCGCCGGGCGTGGCAATCTAGGGGAACGGGACCAAGCTGTCTCTCCAGTG	7300
Db	6025	GCCAAAGTCCCCGGGCGCCGGGCGTGGCAATCTAGGGGAACGGGACCAAGCTGTCTCTCCAGTG	6088
Qy	7301	CACCTGGAGGGAGAAGTGCACAACCGCGGACGCGGCTCAACAACGCGTGTGGAGAGACAG	7366
Db	6085	CACCTGGAGGGAGAAGTGCACAACCGCGGACGCGGCTCAACAACGCGTGTGGAGAGACAG	6144
Qy	7361	CCCTGCTGCGCAGGTTTCAGAGCCATTTCCTCTACAACCCCTGATCATATGCGGCTGCAGGC	7420
Db	6145	CCCTGCTGCGCAGGTTTCAGAGCCATTTCCTCTACAACCCCTGATCATATGCGGCTGCAGGC	6204
Qy	7421	GGATGTCATGGCTTCCCCACCCCCCAACGGAGGCTCCCGCGGGGACAGGGGCCCTGCGTGGC	7480
Db	6205	GGATGTCATGGCTTCCCCACCCCCCAACGGAGGCTCCCGCGGGGACAGGGGCCCTGCGTGGC	6266
Qy	7481	CCCCACCAAGCTTGGAGACGAGAGACCCCAAGCCCATCTGTCTGCTGCAGTACGAGACATC	7540
Db	6265	CCCCACCAAGCTTGGAGACGAGAGACCCCAAGCCCATCTGTCTGCTGCAGTACGAGACATC	6324
Qy	7541	TCCGACAGCGAGTGA 7555	
Db	6325	TCCGACAGCGAGTGA 6339	

	RESULT	11.
	BC0004326	
LOCUS	BC0004326	2842 bp mRNA linear PRI 12-JUL-2001
DEFINITION	Homo sapiens, clone IMAGE:3629785, mRNA, partial cds.	
ACCESSION	BC0004326	
VERSION	BC0004326.1 GI:13279241	
KEYWORDS	.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
REFERENCE	Bukacynski; Metzosa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;	
TITLE	Hominidae; Homo.	
JOURNAL	1 (bases 1 to 2842) Strausberg,R. Direct Submmission Submitted (01-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgabs-r@mail.nih.gov  Tissue Procurement: ATCC  CDNA Library Preparation: Rubin Laboratory DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@cgsc.bc.ca  Susan Jones, Jennifer Asano, Ian Bosdet, Yaron Buterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Kan Guin,	

Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeary, Steven Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline Serier, Diane Smalrus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Teai, Nataasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILM, at: <http://image.llnl.gov>  
Series: IRLM, Place: 13 Row: 1 Column: 14  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES  
source

1. 2842  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3629785"  
/tissue\_type="Placenta, choriocarcinoma"  
/clone\_lib="NIH MGC\_21"  
/lab\_host="DH10B-R"  
/note="vector: pOTB7"  
<1..1821  
/codon\_start=1  
/product="Unknown (protein for IMAGE:3629785)"  
/protein\_id="AAH04326.1"  
/db\_xref="GI:13279242"  
/translation="OAGVWASPLGGTLDGVYPTIMEVLLPKKAPRYAPBPBPADT  
GHAFLAKPPASGLBEPASPSKGSBPRPLVPPVSGHATITAPKALPHASPPPA  
PPASASDPRKRTQSPSISQLELRSLGTHSSYSPBVEVSSPSLTHDGLP  
KHLELDKSHLEGBLRKQPGPVKLGSAHALPHLRPBPESQPSPLQTPAGVKG  
ORVTLADHISEVITODYTRHNPQOLSAPLPAPLYSPGASCPLYDLRPPSDLYLP  
PDHGAAPASGSESGKSPBNPKTSVGGGEGDGIERSPPGMEPBGHSAAVYPL  
YRDEQRTSPSMGTSRGNTSOPRAPEKLTESNAGAKSKKOENKLNTHNREPE  
YNIQSPGTEIRNMPAITGTGLMTRYSQAVQEHASTNMGLEALIRKALMGKTDQNEBP  
PLSAAFPPLNASALPLAMPITLADGSDHTLSPGGGAKAKVSGRSSKASAP  
GLASGDRPSPSVSHSBDGNKRTPLTRVWEDRSPFAGSTPPFPNPLIMLOAGVMA  
SPPPGLPAGSGPLAGPHHAWDEBPRLCSQYETLSDSE"

CDS

Query Match 32.2%; Score 2759.8; DB 8; Length 2842;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 2798; Conservative 2; Mismatches 15; Indels 12; Gaps 2;

5746	TGCCACCCACAGCTGGGGGAGCACTCGATGGGGGTCTACCTTACCTCATGAGGCC	5805
12	TGTCTGGCTTCCCACTGGGCGGCACTCGATGGGGGTCTACCTTACCTCATGAGGCC	71
5806	CGTCTTGTGCCAAGAGAGGCCCCCGGGTCCGCCGCAAGACGCGCCCGAGCAGACAC	5865
72	CGTCTTGTGCCAAGAGAGGCCCCCGGGTCCGCCGCAAGACGCGCCCGAGCAGACAC	131
5866	CGGCGATGCTTCTCTGCCAAGGCCCCCGGCTCCGGGCTGAGAGCCGCTCTCCCG	5925
132	CGGCGATGCTTCTCTGCCAAGGCCCCCGGCTCCGGGCTGAGAGCCGCTCTCTCCCG	191
5926	CAGCAAGGGCTCGAGGCCCGGCGCTGATGCTCTGTCTGTGCAAGCCATGCG	5985
192	CAGCAAGGGCTCGAGGCCCGGCGCTGATGCTCTGTGTGCAAGCCATGCG	251
5986	CGGCAAGGGCTCGAGGGCTCGAGCTTCAACGAGCCGAGCCCGGCGCGCCGAC	6045
252	CGGCAAGGGCTCGAGGGCTCGAGCTTCAACGAGCCGAGCCCGGCGCGCCGAC	311
6046	TGCTCTGGCTCGAGAGCCGCGACCGGAGAAAGCTCAAGTAAACCTTTTCATCAGGA	6105
312	TGCTCTGGCTCGAGAGCCGCGACCGGAGAAAGCTCAAGTAAACCTTTTCATCAGGA	371
6106	ACTGGAAGCTCGGTTCTGGGTTACACAGGCGAGAGCTACAGCCCGGAGAGGGGTGAGGC	6165
372	ACTGGAAGCTCGGTTCTGGGTTACACAGGCGAGAGCTACAGCCCGGAGAGGGGTGAGGC	431

QY	6166	CGTCAGCCCTGTGAGCTCACCCAGTGTGACCCACAGCAGAGGAGGCTCTCCCAAGACCTTGA	6225
DB	432	CGTCAGCCCTGTGAGCTCACCCAGTGTGACCCACAGCAGAGGAGGCTCTCCCAAGACCTTGA	491
QY	6226	AGAGCTGCAAGAGGACCTGTGAGGAGGAGCTGTGGGCCCAAGAGCCAGGAGCCCGTGA	6285
DB	492	AGAGCTGCAAGAGGACCTGTGAGGAGGAGCTGTGGGCCCAAGAGCCAGGAGCCCGTGA	551
QY	6286	GCTTGGCGGAGAGGCGGCCACTTCCACACTGCGGCGGCTGTGAGAGCCAGCCCTC	6345
DB	552	GCTTGGCGGAGAGGCGGCCACTTCCACACTGCGGCGGCTGTGAGAGCCAGCCCTC	611
QY	6346	GTCAGCCCGCTGCTCCAGACCGGCCAGGGGTCAAGAGTCAACGCGGTGTCACTCCT	6405
DB	612	GTCAGCCCGCTGCTCCAGACCGGCCAGGGGTCAAGAGTCAACGCGGTGTCACTCCT	671
QY	6406	GGCCAGAGACATTCATGAGGTTCATCAACAGAGCTACCGGGACCCAGCCAGCAGACT	6465
DB	672	GGCCAGAGACATTCATGAGGTTCATCAACAGAGCTACCGGGACCCAGCCAGCAGACT	731
QY	6466	CAGGCAACCCCTGCGCGGCCCTCTACTCTCTGCGGCGCAGCTGCCCCGTCTGA	6525
DB	732	CAGGCAACCCCTGCGCGGCCCTCTACTCTCTGCGGCGCAGCTGCCCCGTCTGA	791
QY	6526	CCTCGCGGCCACCCAGTGAATCTTACTCTCCGCCCCCGGACCAAGTGCCCGCGCCG	6585
DB	792	CCTCGCGGCCACCCAGTGAATCTTACTCTCCGCCCCCGGACCAAGTGCGCCGCGCCG	851
QY	6586	TGGCTCCCCCAACAGGAGGGGCAAGAGGTCTCCAGGCAACAGAGTGGGTCTT	6645
DB	852	TGGCTCCCCCAACAGGAGGGGCAAGAGGTCTCCAGGCAACAGAGTGGGTCTT	911
QY	6646	GGGTGTGTGAGGACGATATGAACTGTGTCCCAACCGAGGAGCATGACGAGCCAG	6705
DB	912	GGGTGTGTGAGGACGATATGAACTGTGTCCCAACCGAGGAGCATGACGAGCCAG	971
QY	6706	GCACTCCCGAGGTGTGTACCGGCTGTGATCCGGAGTGGGAAACAGAGAGCCAG	6765
DB	972	GCACTCCCGAGGTGTGTACCGGCTGTGATCCGGAGTGGGAAACAGAGAGCCAG	1031
QY	6766	CAGATGGGCTCCAGTCTCAGGCAACAGAGCCGCGCAAGCTCTTCAGCAAGCT	6825
DB	1032	CAGATGGGCTCCAGTCTCAGGCAACAGAGCCGCGCAAGCTCTTCAGCAAGCT	1091
QY	6826	GACCGAGCAACTCCGCGATGTCAAGTCCAGAGCAAGAGATCAACAGAACTGA	6885
DB	1092	GACCGAGCAACTCCGCGATGTCAAGTCCAGAGCAAGAGATCAACAGAACTGA	1151
QY	6886	CACCCACACCGGAATGAGCTTAATATCAATATCAAGCAGCCTGGAGTGTTCAA	6945
DB	1152	CACCCACACCGGAATGAGCTTAATATCAATATCAAGCAGCCTGGAGTGTTCAA	1211
QY	6946	TATGCGGCGCATCAACCGGAACAGGCTTAATAGACCTATGAAGCCAGGGGTGAGAA	7005
DB	1212	TATGCGGCGCATCAACCGGAACAGGCTTAATAGACCTATGAAGCCAGGGGTGAGAA	1271
QY	7006	TGCCAGCAACAATGGGGCTGAGAGCCATATTAAGAAAGCACTCATGGTAAATATGA	7065
DB	1272	TGCCAGCAACAATGGGGCTGAGAGCCATATTAAGAAAGCACTCATGGTAAATATGA	1331
QY	7066	CGATGGGAAGATCCCGCGCTCAAGCCGCAATGCTTTAAACCTCTGAATGCCAGTGC	7125
DB	1332	CGATGGGAAGATCCCGCGCTCAAGCCGCAATGCTTTAAACCTCTGAATGCCAGTGC	1391
QY	7126	CAGCTGCGCGCTGATAGCCATTAACCGGTGAGAGGAGAGTGAACACACTGAC	7185
DB	1392	CAGCTGCGCGCTGATAGCCATTAACCGGTGAGAGGAGAGTGAACACACTGAC	1451
QY	7186	CTGCGCAGGTGGCGGAGGAGCCAAAGTCTCTGGCAGAACCCAGCAGCCGAAAGCCAA	7245
DB	1452	CTGCGCAGGTGGCGGAGGAGCCAAAGTCTCTGGCAGAACCCAGCAGCCGAAAGCCAA	1511
QY	7246	GTCGCCGAGCCCGGCGCTGAGATCTGGGAGCCGAGCACTCTGTCTCTCATGAGCACTC	7305

Db 1512 GTCCCGGCGCCCGGCGCTGGCATCTGGGAGACCGGCGCACCTCTGTCTCTCAATGACATC 1571  
Qy 7306 GAGAGGAGACTGCAACCGCGCGAGCGCGCTCAACCAACCGGATGTGGAGAGAGAGCGCCCTC 7365  
Db 1572 GAGAGGAGACTGCAACCGCGCGAGCGCGCTCAACCAACCGGATGTGGAGAGAGAGCGCCCTC 1631  
Qy 7366 GTCCGAGGTTCCACCGCATCTCCCTCAACACCCCTGATCATCGGCTGACGCGGGTGT 7425  
Db 1632 GTTCGAGGTTCCACCGCATCTCCCTCAACACCCCTGATCATCGGCTGACGCGGGTGT 1691  
Qy 7426 CATGCTTCCCGACCCCGACCGGCGCTCCCGCGGAGCGGCGCCCTGCTGCGCCCGCA 7485  
Db 1692 CATGCTTCCCGACCCCGACCGGCGCTCCCGCGGAGCGGCGCCCTGCTGCGCCCGCA 1751  
Qy 7486 CCAAGCGCTGGAGAGAGAGAGCGCCACGCACTGCTGTGTGAGATAGAGACATCTCCGA 7545  
Db 1752 CCAAGCGCTGGAGAGAGAGAGCGCCACGCACTGCTGTGTGAGATAGAGACATCTCCGA 1811  
Qy 7546 CAGCGAGTACTCAGAACAGAGCGGCGGAGGCGGCGGCTGEBRSPRLUGTCAGGTC 7605  
Db 1812 CAGCGAGTACTCAGAACAGAGCGGCGGAGGCGGCGGCTGEBRSPRLUGTCAGGTC 1860  
Qy 7606 ACGAGAGCAACAGAGAGCGGCGCTGAGAGAGCGGCGGCTGCGCACTCCCGCAACCAAGA 7665  
Db 1861 ACGAGAGCAACAGAGAGCGGCGCTGAGAGAGCGGCGGCTGCGCACTCCCGCAACCAAGA 1920  
Qy 7666 AAGAGCGGCTGAGTCCGCGCTGCGCGCTCCATCCATCTGTCTGTGAGAGCGGAGATCTTG 7725  
Db 1921 AAGAGCGGCTGAGTCCGCGCTGCGCGCTCCATCCATCTGTCTGTGAGAGCGGAGATCTTG 1980  
Qy 7726 CCGTGTAAAGCGCTTAAGACTCCCGCGCGGCGGCGCTGAGAGCGGAGATCTTG 7785  
Db 1981 CCGTGTAAAGCGCTTAAGACTCCCGCGCGGCGGCGCTGAGAGCGGAGATCTTG 2040  
Qy 7786 GGGAGATGTTACTGTGTCTGCGGAGAGAGAGGAGAGGCGCGGAGAGAGGCGGAG 7845  
Db 2041 GGGAGATGTTACTGTGTCTGCGGAGAGAGAGGAGAGGCGCGGAGAGAGGCGGAG 2100  
Qy 7846 GCGTGTGAGAGCAACAGAGCGGCGGAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAG 7905  
Db 2101 GCGTGTGAGAGCAACAGAGCGGCGGAGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAG 2160  
Qy 7906 GCACTCTCAAGCGCACTGCGCTCCCGCGAGATGATTTGAGAACCAAGCTTAACTGAGCTG 7965  
Db 2161 GCACTCTCAAGCGCACTGCGCTCCCGCGAGATGATTTGAGAACCAAGCTTAACTGAGCTG 2220  
Qy 7966 CAGCGCGCGCGCGCTCCCTCCCGCTCCATCCCGCTTAAAGCGCTTGAAGAGATGAGAGCA 8025  
Db 2221 CAGCGCGCGCGCGCTCCCTCCCGCTCCATCCCGCTTAAAGCGCTTGAAGAGATGAGAGCA 2280  
Qy 8026 GCGCGCTGTCAAGCGCGCGCGAGTGGCTGTTCGGTCCCGCAAGAGCTGCCCGCAAGAG 8085  
Db 2281 GCGCGCTGTCAAGCGCGCGCGAGTGGCTGTTCGGTCCCGCAAGAGCTGCCCGCAAGAG 2340  
Qy 8086 ATTGCTGAGAAACCAAGTCAAGCGAGGTGGAGCGGAGCAAAAGGCGAGGTGGCGCTGGGG 8145  
Db 2341 ATTGCTGAGAAACCAAGTCAAGCGAGGTGGAGCGGAGCAAAAGGCGAGGTGGCGCTGGGG 2400  
Qy 8146 GAAAGGATGCTCCGAGAGCTGAGCTGTTTTTTCACACATCGTTGCGGAGCGGAGGAGAA 8205  
Db 2401 GAAAGGATGCTCCGAGAGCTGAGCTGTTTTTTCACACATCGTTGCGGAGCGGAGGAGAA 2460  
Qy 8206 GGAAGGAGAGATGATTAATATGTGTGTTTACAGGGATATATTTTGAATACCTTCAATGA 8265  
Db 2461 GGAAGGAGAGATGATTAATATGTGTGTTTACAGGGATATATTTTGAATACCTTCAATGA 2520  
Qy 8266 ATTAATTCAGATGTTTACGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8325  
Db 2521 ATTAATTCAGATGTTTACGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580  
Qy 8326 GATCTCTGCTTACCGTTCAAGAGCGGTGTGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAG 8385  
Db 8326 GATCTCTGCTTACCGTTCAAGAGCGGTGTGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAG

Db 2581 GATCTCTGCTTACCGTTCAAGAGCGGTGTGAGGCGGAGAGAGAGAGAGAGAGAGAGAGAG 2640  
Qy 8386 CAGAGCAAG 8444  
Db 2641 CAGAGCAAG 2700  
Qy 8445 CTTTGGGAGAGATGAAATTCGATGCGTATCTGTGCGCGAGATTTTGGCGAGGAGTGTGTA 8504  
Db 2701 CTTTGGGAGAGATGAAATTCGATGCGTATCTGTGCGCGAGATTTTGGCGAGGAGTGTGTA 2760  
Qy 8505 TTCTGTATTTTACACAGCTGCTGTTTAAATTAAGAGATTAATTAATTAATTAATTAATTA 8564  
Db 2761 TTCTGTATTTTACACAGCTGCTGTTTAAATTAAGAGATTAATTAATTAATTAATTAATTA 2820  
Qy 8565 AAAAAA 8571  
Db 2821 AAAAAA 2827

RESULT 12  
S83390  
LOCUS 2930 bp mRNA linear PRI 28-MAR-1997  
DEFINITION T3 receptor-associating cofactor-1 [human, fetal liver, mRNA, 2930 nt].  
ACCESSION S83390  
VERSION S83390.1 GI:1911769  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 2930)  
Sande, S. and Privalsky, M.L. Identification of TRACs (T3 receptor-associating cofactors), a family of cofactors that associate with, and modulate the activity of, nuclear hormone receptors  
Mol. Endocrinol. 10 (7), 813-825 (1996)  
JOURNAL PubMed 8813722  
REMARK Genbank staff at the National Library of Medicine created this entry [NCBI glibbeq 179975] from the original journal article.  
FEATURES  
source  
1..2930  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..2930  
/gene="T3 receptor-associating cofactor-1, TRAC-1"  
456..2765  
/gene="T3 receptor-associating cofactor-1, TRAC-1"  
/note="Authors report detecting several TRAC alternative transcripts; TRAC-1"  
/codon\_start=1  
/product="T3 receptor-associating cofactor-1"  
/protein\_id="AAB50847.1"  
/db\_xref="GI:1911770"  
/translation="MAQPADMRGLSPRESSALNYAAGPGIIDLQVPHLVYVP  
TGTPATMDRLAYLPTAPQPPSRHSSSPSPGCTHTTKTTSSSRBRDRBR  
DREDEREKXSLTSTTVHEAPLWRPTEGSSSSGSSGGGSSSPASHAHQHP  
ISPRQDALQORPSVLHNTGMKGIITAVBPSTVLRSTSSPVAPATFPATCP  
LGTLGDVPTLMEVLPKEAPRVARPERPADTGATLAKPAPAGLEPASPSPK  
SEPRPLVPVSGHATARTPAKNLAPHNAPPPAPAPASDPHREKTSKPSIQEL  
ELRSIGYSSSPGCVREPVSPSLTHDKGLPHNLELHLSHLEGLRPPQRPV  
KLGEERAHLPRLPSPSSPPLQTPAPGVKHORVYTLAQHISEVYTDYTRHP  
QQLSAPLPAVLSPFGASCPLVDLRPBDLVLPDRDAPAGSPHSGKSPSPN  
KTSVLGGGEGDGLPVPSPGKTEPGHRSRAVPLVLRDQEPSPHSGKSPSPN  
PAFSPKLTESNANVSKKOEINKLNTNRMEPEKNISQPTBLIFNMAIIGTGLMT  
YRSQAVOEASTNMGLBALIRKALMGGAGKAVSGSPSRKAKSPAPGASGDRPSPV  
SSVHSGCNRPTPLTNVWEDRPSASSTPPPNVPLIWRLDAQWASPPPGGLPAGS  
GPLAGNHAHNDREPRPLCSQYETISDS"

ORIGIN  
Query Match 30.6%; Score 2618.6; DB 8; Length 2930;



Best Local Similarity 94.6%; Pred. No. 0; Matches 2786; Conservative 3; Mismatches 9; Indels 146; Gaps 2;									
QY	4785	CCCCGACAGCAGCCGTCGCCGAGACCAACCCACACCCCATCTCGCCCTATGAGCACTGCG	4844						
Db	133	CCCCGACAGCAGCCGTCGCCGAGACCAACCCACACCCCATCTCGCCCTATGAGCACTGCG	192						
QY	4845	TTTCGGGGGTGAGTGGCTGTGACCTGTATGAGGCAACATCCGCCCTTGACCTTGAACCCCA	4904						
Db	193	TTTCGGGGGTGAGTGGCTGTGACCTGTATGAGGCAACATCCGCCCTTGACCTTGAACCCCA	252						
QY	4905	CCTCCATACCCCGCGGCATCTCTTGAGAGCAGCGGCTGCTTACTTACTTCTGCCCCGACACC	4964						
Db	253	CCTCCATACCCCGCGGCATCTCTTGAGAGCAGCGGCTGCTTACTTACTTCTGCCCCGACACC	312						
QY	4965	TGGCCCCCAACCCCACTTACCCGCACTGTACCCACCCCTACCTCATCCGCGGTACCCCG	5024						
Db	313	TGGCCCCCAACCCCACTTACCCGCACTGTACCCACCCCTACCTCATCCGCGGTACCCCG	372						
QY	5025	ACACGGCGCGCTGAGAAACCGGCAAGACCATCATCATGACTACATCTGCGACAGACA	5084						
Db	373	ACACGGCGCGCTGAGAAACCGGCAAGACCATCATCATGACTACATCTGCGACAGACA	432						
QY	5085	TGCAACAACAACGGGCACCGCCATGAGCCCAAGAGCTGATATGCTGAGGGGCTCTCGC	5144						
Db	433	TGCAACAACAACGGGCACCGCCATGAGCCCAAGAGCTGATATGCTGAGGGGCTCTCGC	492						
QY	5145	CCCGGAGTCTCGTGGGCACTCAACTAGCTGCGGGGTCCCGGAGCATCATGACCTGT	5204						
Db	493	CCCGGAGTCTCTGCTGGGCACTCAACTAGCTGCGGGGTCCCGGAGCATCATGACCTGT	552						
QY	5205	CCCAAGTGCAACCTGCTGTGTCTGTGCCCCCGACACACAGGCAACCCAGCCACCGCCA	5264						
Db	553	CCCAAGTGCAACCTGCTGTGTCTGTGCCCCCGACACACAGGCAACCCAGCCACCGCCA	612						
QY	5265	TGGACCGCTTGGCTTACTTCTCCCAACCGGCCCCCAAGCTTTACAGACCGCCCAAGAGCT	5324						
Db	613	TGGACCGCTTGGCTTACTTCTCCCAACCGGCCCCCAAGCTTTACAGACCGCCCAAGAGCT	672						
QY	5325	CCCACTCTCCCAAGAGGTCCAAACAATTGACAAAACAACAACAACAGTCTGTGCG	5384						
Db	673	CCCACTCTCTCCCAAGAGGTCCAAACAATTGACAAAACAACAACAACAGTCTGTGCG	732						
QY	5385	AGCGGAGGAGACCGGAGTCGAGAGCGGGAACCGGATCGGAGCGGGAAGATCCATCC	5444						
Db	723	AGCGGAGGAGACCGGAGTCGAGAGCGGGAACCGGATCGGAGCGGGAAGATCCATCC	792						
QY	5445	TCAAGTCCACACACGATGAGACACGACCCATCTGAGAACCTGTGTACAGACAGACGA	5504						
Db	793	TCAAGTCCACACACGATGAGACACGACCCATCTGAGAACCTGTGTACAGACAGACGA	852						
QY	5505	GGGGAGAGCGGAGCAGCGCGGGGGGTGGGGGACGACAGCGGCCCCGCTCCACT	5564						
Db	853	GGGGAGAGCGGAGCAGCGCGGGGGGTGGGGGACGACAGCGGCCCCGCTCCACT	912						
QY	5565	CCCATGACCACGAGACTTCGCCATCTCCCTCGGACCCAGAGATGCTTCAGACAGAGAC	5624						
Db	913	CCCATGACCACGAGACTTCGCCATCTCCCTCGGACCCAGAGATGCTTCAGACAGAGAC	972						
QY	5625	CCAGTGTGTTTCAACAACAGGCAATGAGGGTATCATCACCGGTGTGAGCCCAAGAC	5684						
Db	973	CCAGTGTGTTTCAACAACAGGCAATGAGGGTATCATCACCGGTGTGAGCCCAAGAC	1032						
QY	5685	CCAGGATCTGAGGTCACTCACTCACTCTCAACCGGTTGAGCCAGTGCACATTTCCAC	5744						
Db	1033	CCAGGATCTGAGGTCACTCACTCACTCTCAACCGGTTGAGCCAGTGCACATTTCCAC	1092						
QY	5745	CTGCAACCACTGACCACTTGGGCGGCAACCTGATGGGGTTTAACTTCACTTCAATGAGC	5804						
Db	1093	CTGCAACCACTGACCACTTGGGCGGCAACCTGATGGGGTTTAACTTCACTTCAATGAGC	1152						
QY	5805	CGGTCTTGTGCCCCAAGAGGCCCGCGGGTGGCCCGGCAGAGCGGCCCGACAGACA	5864						

Db	1153	CGGTCTTGTGCCCCAAGAGGCCCGCGGGTGGCCCGGCAGAGCGGCCCGACAGACA	1212						
QY	5865	CCGACATGCTCTTCTGCGCAAGCCGCCAGCCGCTCGGAGTGTGAGCCGCTCTCC	5924						
Db	1213	CCGACATGCTCTTCTGCGCAAGCCGCCAGCCGCTCGGAGTGTGAGCCGCTCTCTCC	1272						
QY	5925	CCAGCAAGGGCTGGAGGCCCGGCTTATGTCCTTCTGTCTGTGACACGCAACCATG	5984						
Db	1273	CCAGCAAGGGCTGGAGGCCCGGCTTATGTCCTTCTGTCTGTGACACGCAACCATG	1332						
QY	5985	CCCGAACCCCTGCGAAGAACCTTCGACCTCAACAGCCAGCCCGGACCCGCGGCGCAC	6044						
Db	1333	CCCGAACCCCTGCGAAGAACCTTCGACCTCAACAGCCAGCCCGGACCCGCGGCGCAC	1392						
QY	6045	CTGCTCGGCTCGGACCGGCAACCGGAAAGAACTCAAGTAAACCTTTTCATCCAG	6104						
Db	1393	CTGCTCGGCTCGGACCGGCAACCGGAAAGAACTCAAGTAAACCTTTTCATCCAG	1452						
QY	6105	AACTGAACTCGTTCTGTGGTTACACGCGACAGCTTACAGCCCGGAGGGTGTAGAC	6164						
Db	1453	AACTGAACTCGTTCTGTGGTTACACGCGACAGCTTACAGCCCGGAGGGTGTAGAC	1512						
QY	6165	CCGTAGCCCTGTGAGCTTCAACCAATGAGCCAGCAAGAGGGGCTCCCAAGACCTGG	6224						
Db	1513	CCGTAGCCCTGTGAGCTTCAACCAATGAGCCAGCAAGAGGGGCTCCCAAGACCTGG	1572						
QY	6225	AAGACTCGCAAGAAGCACTTGAGAGGGAGCTGCGGCCAAGGACCGGCGCCGTGA	6284						
Db	1573	AAGACTCGCAAGAAGCACTTGAGAGGGAGCTGCGGCCAAGGACCGGCGCCGTGA	1632						
QY	6285	AGCTTGGGGGAGGCGCGCCACCTTCCACACTGCGCGCTGCTGTAGAGCCAGCCT	6344						
Db	1633	AGCTTGGGGGAGGCGCGCCACCTTCCACACTGCGCGCTGCTGTAGAGCCAGCCT	1692						
QY	6345	CGTCCAGCCGCTGCTCCAGAACCGGCCCCGAGGTCAAGGTCAACAGGGGTGATACCC	6404						
Db	1693	CGTCCAGCCGCTGCTCCAGAACCGGCCCCGAGGTCAAGGTCAACAGGGGTGATACCC	1752						
QY	6405	TGGCCAGCACTAGTGAAGTCACTACACAGACTACACCCGCGACACCCACAGCAAC	6464						
Db	1753	TGGCCAGCACTAGTGAAGTCACTACACAGACTACACCCGCGACACCCACAGCAAC	1812						
QY	6465	TGAGCGCACCCCTGCGCCCCCTCTTACTCTTCCCTGAGGCGACGTCGCCCTCTGG	6524						
Db	1813	TGAGCGCACCCCTGCGCCCCCTCTTACTCTTCCCTGAGGCGACGTCGCCCTCTGG	1872						
QY	6525	ACCTCGCGGCGCAACCGATGACTCTACCTCGCGCCCCCGGACCAATGGTGCCTCGGCC	6584						
Db	1873	ACCTCGCGGCGCAACCGATGACTCTACCTCGCGCCCCCGGACCAATGGTGCCTCGGCC	1932						
QY	6585	GTGGCTCCCCCAAGCGAAGGGGCAAGAGGTCTTCAGAGCCAAACAAGACGTGCT	6644						
Db	1933	GTGGCTCCCCCAAGCGAAGGGGCAAGAGGTCTTCAGAGCCAAACAAGACGTGCT	1992						
QY	6645	TGGGTGTGTGAGAGCGGTATTTGAACTGTGTCCCAACCGGAGGGCATGACGAGCCAG	6704						
Db	1993	TGGGTGTGTGAGAGCGGTATTTGAACTGTGTCCCAACCGGAGGGCATGACGAGCCAG	2052						
QY	6705	GCGACTCCCGAGATCTGTGTACCCGCTGTGTACCCGGATGGGAACAAGAGGCCCA	6764						
Db	2053	GCGACTCCCGAGATCTGTGTACCCGCTGTGTACCCGGATGGGAACAAGAGGCCCA	2112						
QY	6765	GAGAGTGGGCTCCAAAGTCTCCAGGCAACACAGCAGCGCCAGCTTCTTCAGCAAC	6824						
Db	2113	GAGAGTGGGCTCCAAAGTCTCCAGGCAACACAGCAGCGCCAGCTTCTTCAGCAAC	2172						
QY	6825	TGACGAGAGCACTCCGCAATGTCTCAAGTCAAGTCAAGAGCAAGATCAACAAGACTGA	6884						
Db	2173	TGACGAGAGCACTCCGCAATGTCTCAAGTCAAGTCAAGAGCAAGATCAACAAGACTGA	2232						
QY	6885	ACAACCAACAACGGGAATGAGCTTGAATTAATATACAGCCAGCTGGGAGAGGATTTCA	6944						
Db	2233	ACAACCAACAACGGGAATGAGCTTGAATTAATATACAGCCAGCTGGGAGAGGATTTCA	2292						

```

QY      6945 ATATGCCCGCCATCACCGGAACAGGCGCTTATGACCTTATAGAACAGGCGGTCAGAGAAC 7004
      |||
      2293 ATATGCCCGCCATCACCGGAACAGGCGCTTATGACCTTATAGAACAGGCGGTCAGAGAAC 2352
QY      7005 ATGCCAGACCAACATGAGGCGTGAAGCCATTAATTAGAAAGGCACTGATGGTAAATATG 7064
      |||
      2353 ATGCCAGACCAACATGAGGCGTGAAGCCATTAATTAGAAAGGCACTGATG----- 2402
QY      7065 AACAGTGGGAAGTCCCGCGCTGAGGCCCAATGCTTTAAACCTCTGAAATGCCAGTG 7124
      |||
      2403 ----- 2402
QY      7125 CCAGCTGCGCGCTGATGATGCCATAACCGCTGTGTACGAGCGAGATGACACACTCA 7184
      |||
      2403 ----- 2402
QY      7185 CCTGCGCAGGTGCGCGCGGAGAGCCAAAGTCTCTGCGAGACCCAGCAGCCGAAAGCCA 7244
      |||
      2403 ----- 2402
QY      7245 AGTCCCGCGCGCGCGCTGAGCATCTGAGGAGACCGGCAACCTCTGTCTCTCACTGCACT 7304
      |||
      2455 AGTCCCGCGCGCGCGCTGAGCATCTGAGGAGACCGGCAACCTCTGTCTCTCACTGCACT 2514
QY      7305 CGGAGGAGAGACTGCAACCGCGGAGACCGCGCTCAACACCGCGTGTGAGAGAGAGAGCCCT 7364
      |||
      2515 CGGAGGAGAGACTGCAACCGCGGAGACCGCGCTCAACACCGCGTGTGAGAGAGAGAGCCCT 2574
QY      7365 CGTCCGAGAGTTTCACGCGCATTTCCCTCAACACCCCTGATCATGCGGCTGCAAGCGGCT 7424
      |||
      2575 CGTCCGAGAGTTTCACGCGCATTTCCCTCAACACCCCTGATCATGCGGCTGCAAGCGGCT 2634
QY      7425 TCATGCGTTCCCAACCGCGCGCTCCCGCGGAGCGGCGCGCTGCTGCGCGCC 7484
      |||
      2635 TCATGCGTTCCCAACCGCGCGCTCCCGCGGAGCGGCGCGCTGCTGCGCGCGCC 2694
QY      7485 ACCAGCGCTGGGAGCGAGAGCCCAAGCACTGCTGCTGCGAGTACGAGACACTCTCCG 7544
      |||
      2695 ACCAGCGCTGGGAGCGAGAGCCCAAGCACTGCTGCTGCGAGTACGAGACACTCTCCG 2754
QY      7545 ACAGCAGTGACTCAGAACAGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 7604
      |||
      2755 ACAGCAGTGACTCAGAACAGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2806
QY      7605 CAGCGAGCGACAGGAAACCGGCGCTGCGAGAGCGGCGGCGCTGCTGCTGCTGCTGCTGCT 7664
      |||
      2807 CAGCGAGCGACAGGAAACCGGCGCTGCGAGAGCGGCGGCGCTGCTGCTGCTGCTGCTGCT 2866
QY      7665 AAGGAGCCCTGAGTCCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7724
      |||
      2867 AAGGAGCCCTGAGTCCGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2926
QY      7725 GCCT 7728
      |||
      2927 GCCT 2930

```

```

RESULT 13
LOCUS   BC047524          2964 bp      mRNA      linear      ROD 05-NOV-2003
DEFINITION Mus musculus nuclear receptor co-repressor 2, mRNA (cDNA clone
ACCESSION BC047524
VERSION   BC047524.1 GI:29126781
KEYWORDS
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
          Sciurognathi; Murioidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2964)
AUTHORS  Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

```

```

FEATURES
source
gene
CDS

```

Klausener, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buelow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marinina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Useth, T.B., Toshynuk, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mulhally, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 2964)  
 Strausberg, R.  
 Direct Submission  
 Submitted (03-MAR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: The Cepko Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
 Center, Stanford University School of Medicine, Stanford, CA 94305  
 Web site: <http://www-shgc.stanford.edu>  
 Contact: (Dickson, Mark) [mcdpaxi.stanford.edu](mailto:mcdpaxi.stanford.edu)  
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>  
 Series: IRAX Series: 44 Row: 1 Column: 22  
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

Location/Qualifiers  
 1..2964  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4526899"  
 /issue\_type="Eye, retina, mouse strain C57BL/6"  
 /clone\_id="NIH\_MGC\_94"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORTe"  
 <1..2964  
 /gene="Ncor2"  
 /note="synonyms: SMRT, SMRTe, N-COR"  
 /db\_xref="GeneID:20602"  
 /db\_xref="MGI:1337080"  
 <1..1896  
 /gene="Ncor2"  
 /codon\_start=1  
 /product="Ncor2 protein"  
 /protein\_id="AAH47524.1"  
 /db\_xref="GI:29126782"  
 /db\_xref="GeneID:20602"  
 /db\_xref="MGI:1337080"  
 /translation="VTSVEPGPTVLRWARSSTSSPVRAATPPATCPGLTLEGVPTLMRPVTLPKETSRVARPERPRVDAGHAFLTKEPPAEPASSSSKSPSLAPS SHTAIARTPAKLAIPHASPDPAPTSADLHREKTSKPPSIQELERISLGYHSGA

ORIGIN

GYSPDVEIPISVSPSSPLTHDKLSKPLEBELEKSHLBELRHKQBPENKLSABANAHP  
HLRPLBESQPSPPSLLQTPGIKQHVVTSAKOHISEVITDQYTHNHQOOLSGPLAP  
LVSFPGASCPLVLDLRPPSDLYLPDPDTPARSPBEGKRSPEPSKTSVLSGSD  
AIEPVSPEGMTEPGRHASTAYPLIYRQBOEPBMGSKSPNTSOPAPSPKLTRESN  
SAMYKSKQBJENKIKNTNRNRPENPEYINQPGREIIPMRAITGAGMTCSRQAVQDHS  
TNMGJEAIIIRALMKGIYQWEPPLVANAFLPLNASLPLPAAPITTAIDRSJHAL  
TSFGGGGAKVSGRPSRKSAPGLASGDPPSVSGSVLASHGNDNRRTPLNTRWED  
RPSAGSTYFPFYNPLIMRLQAGVNASPPPGILAAGSGLAGHHAMDEBPXPLCSQY  
ETLSDSB"

Query Match 19.9%; Score 1703.4; DB 9; Length 2964;  
Best Local Similarity 79.1%; Pred. No. 5.8e-212;  
Matches 2347; Conservative 1; Mismatches 511; Indels 108; Gaps 23;

QY 5661 TCACCGCTGTGGAGGCCAGCAAGCCACGAGTCTG-----AGGTCACCTCCACT 5711  
DB 2 TCACCTCGGTGAAACCCGGCACGCCACGATCTTGAAGTGGGCGCAGGTCCACTCCACT 61  
QY 5712 CCTCACCCGTTGCCCCAGCTGCACATTCCACCTGCCACCCACCTGCCACTGGCGGCA 5771  
DB 62 CTGGGCTGTGCGGCCAGCTGCACATTCCACCTGCCACCCACCTGCCACTGGGAGCA 121  
QY 5772 CCCTGATGGGGTCTAACCTTACCTCATGAGCCCGTCTTGCTGCCCAAGAGGCCCCC 5831  
DB 122 CCCTTGAAGGGGTCTTACCTTACCTCATGAGCCCGTCTTCTTACCCCAAGAGACTCTTC 181  
QY 5832 GGGTGGCCCGGCGAGAGGGGCCCCGAGAGACACCGGCGCATGCTTCCCTGGCCAGGCC 5891  
DB 182 GGGTGGCCCGGCGAGAGGGGCCCCGATGGAGCGCTGGCCATGCTTCTTCAACCAACCCC 241  
QY 5892 CAGCCGCTCCGGGGTGAAGCCCGGCTCTCCCGCAGCAAGGGCTTCGAGGCCCGCGCC 5951  
DB 242 CGGGCCG-----GGAGCCCGGCTCTCTTCAACCCAGCAAGAGCTCCAGGCCCGATCCC 292  
QY 5952 TAGTGCCTCTGTCTGTGGCCAGCGCCACCATGCGCGCACCCCTTGGAAGAACTTCGAC 6011  
DB 293 TAGCACCCGCCGCTTCAGCGCACACAGCCATGCGCGCACCCCAAGAAAGAACTTGGAC 352  
QY 6012 CTACACAGCGCGCGGAGCCCGGCGGCGGCGACCTGCGCTGGGCTTCGAGCCCGGACCG 6071  
DB 353 CCACCATGCGCAGTCCGAGCCCGCGGCG-----CCACCTGGGCTTCAGATTCGACCGAG 409  
QY 6072 AAAAGACTCAAGTAACCCCTTTTCCATCAGAGAACTGGAACTCCGTTCTCGGGTTACC 6131  
DB 410 AAAAGACTCAAGTAACCCCTTTTCCATCAGAGAACTGGAACTCCGTTCTCGGGTTACC 469  
QY 6132 AC---GGCAGCAGCTTACAGCCCCGAAAGGGGTGAGCCCTCAGCCCTGTGAGCTCAACCA 6188  
DB 470 ACAGTGGAGCTGTGCTTACAGCCCCGATGGGGTGAAGCCCATGAGCCCGGTGAGCTCCCCA 529  
QY 6189 GTCTGACCCAGCAAGAGGGGCTCCCAAGACATCTGGAAAGCTTCGACAAAGACCACTGG 6248  
DB 530 GCTTACCCAGCAAGAGGGGCTCTCAAACTCTTGAAAGACTTGAAGAAAGACCACTTGG 589  
QY 6249 AGGGGAGCTGCGGCCCAAGCAGCAAGGCCCGCTGAGAGCTTGCGGGAGAGCGCGCCACC 6308  
DB 590 AAGGGAGCTGCGGCCCAAGCAGCAAGGCCCGCATGAGCTTCAAGCGGGAGGCTGGCCATTC 649  
QY 6309 TCCCACTCTGCGGCCCACTGCGCGAGCAGCCCTCATCTAGCCCACTCTTCAGACTG 709  
DB 650 TCCCACTCTGCGGCCCACTGCGCGAGCAGCCCTCATCTAGCCCACTCTTCAGACTG 709  
QY 6369 CCCCAAGGGGTCAAAAGGTCAACAGGGGGTGTCAACCTGCGCCAGACATCAGTGAAGTCA 6428  
DB 710 CCCCAAGGCAATCAAAAGGTCAACAGAGGGTGTCACTCGGCTAGACATCAATGAGTGA 769  
QY 6429 TCACACAGGACTTACACCGCGCACCAACCAAGACAGCTTACGCGCACCCCTGCGCGCCCC 6488  
DB 770 TTAACGAGGACTTACACGCGCCACCAACCGCGCACAGTCTAGTGGCCCTTCCGCGCCCTC 829  
QY 6489 TCTACTCTTCTCTTGGGAGCAGCTGCCCCGTCTGTGAACTTCGCGGCGCAACCAAGTACC 6548  
DB 6548 TCTACTCTTCTCTTGGGAGCAGCTGCCCCGTCTGTGAACTTCGCGGCGCAACCAAGTACC 6548

DB 830 TCTACTCTTCTCTTGGGAGCAGCTGCCCCGTCTGTGAACTTCGCGGCGCAACCAAGTACC 889  
QY 6549 TCTACTCTCTCCCGCCCGGACCATGTGTCCCGGCGCGCTGAGCTCCCCCAAGCAAGAGGCG 6608  
DB 890 TCTACTCTCCACCCCGGACCATGTGCACACCCGCGGGGATCCCCCAAGTGAAGGGGG 949  
QY 6609 GCAAGAGTCTTCAGAGCCAAACAAGACGTGTGTTGGGTGTGTGTGAAGACGTTATTG 6668  
DB 950 GCAAAAGGTCCCGAGAACCCAGCAAAACATGTGTCTGTGGGAGCGACAGAGATGCAATTG 1009  
QY 6669 AACCTGTCTCCCAACCGGAGGGCAATGAGGAGCCAGGGGCACTCCGAGTGTGTATCC 6728  
DB 1010 AGCTGTGTCTCCCAACAGAGGGCAATGACTGAGCCAGACATGTCTGAGCACTGTATCC 1069  
QY 6729 CGTGTGTATCCCGGATGGGAAACAGACGAGCCGAGAGATGGGCTTCAAGTCTCAG 6788  
DB 1070 CACTGTGTATGTGAAGCGGGGAACAGGGCGAGCC---CAGAGTGGGCTCTAAGTCTCAG 1126  
QY 6789 GCAACACAGCCAGCGCCAGCTTCTTCAAGACTGACCGAGACAACTTCGCGCATGG 6848  
DB 1127 GCAACACAGCCAGCGCCAGCTTCTTCAAGACTGACCGAGACAACTTCGCGCATGG 1186  
QY 6849 TCAAGTCCAGAGGAGAGATCAACAGAAAGTGAACACCCACACCGGAATGAGCTTG 6908  
DB 1187 TGAAGTCAAGAGAGAGAGATCAACAGAAACTCAACACCAACCGGAACGAGCAG 1246  
QY 6909 AATACATATGACGCGCTGAGACGAGATCTTCAATATGCGCGCATCAACCGGAACAG 6968  
DB 1247 AATACATATGAGCGCGCTGAGACGAGATCTTCAATATGCGCGCATCAACCGGAACAG 1306  
QY 6969 GCTTATGACCTATGAGCCAGCGGTGTGAGAACATGCCAGCACCAATGGGGCTGG 7028  
DB 1307 GCTTATGACCTGTGAGAGCAGGGGTGTGAGAAACAGCGCACCAATGGGGCTAG 1366  
QY 7029 AGGCAATATTAAGAAAGGACATCATGGGTAATATGACCAAGTGGGAGAGTCCCGCGCG 7088  
DB 1367 AGGCAATATTAAGAAAGGACATCATGGGTAATATGACCAAGTGGGAGAGTCCCGCGCG 1426  
QY 7089 TCAGCGCAATGCTTTTAACTCTGAATGCGCAGTGCAGCGCTGCC---GCTGTATATG 7145  
DB 1427 TCAGCGCAATGCTTTTAACTCTGAATGCGCAGTGCAGCGCTGCC---GCTGTATATG 1486  
QY 7146 CCATTAACGCTGTGACGAGCGAGTGAACACACACTCACTTCGCGAGTGGCGGGA 7205  
DB 1487 CCATTAACGCTGTGACGAGCGAGTGAACACGCACTCACTTCGCGAGTGGCGGGA 1546  
QY 7206 AGGCAAGGTCTGTGGAGACCCAGCAGCCGAAAGCAAGTCCCGGCGCGGCGCTGG 7265  
DB 1547 AAGCAAGGTCTGTGGAGACCCAGCAGCCGAAAGCAAGTCCCGGCGCGGCGCTGG 1606  
QY 7266 CATGTGGGAGCCGAGCACTCTGTCTCTCAAGTGCATCTGGAGGAGACTGCAACGCG 7325  
DB 1607 CATGTGGGAGCCGAGCACTCTGTCTCTCAAGTGCATCTGGAGGAGACTGCAACGCG 1666  
QY 7326 GAGCGCGCTCAACACGCGGTGTGGAGAGACAGCCCTGTTCGCAAGTTTCAAGCCAT 7385  
DB 1667 GAAACACACTCAACCAACGCTGTGTGGAGAGCCGCGCTCATCTGCAAGGTTCAAGCCAT 1726  
QY 7386 TCCCTTAACACCCCGTGAATCATGGGCTGAGGGGGTGTCAATGACTTCCCAACCCCAAC 7445  
DB 1727 TCCCTTAACACCCCGTGAATTAAGGCTTACAGGAGGTCAATGAGCTTCCCGCCCAAC 1786  
QY 7446 CGGAGCTCCCGCGGAGCGGAGCCCTGTGTGAGCCCGCCACCAACGCTTGGAGAGAGAC 7505  
DB 1787 CTGGGCTTGGCGGAGGAGCGGAGCCCTTGTGTGTGAGCTGAGCTTGGAGAGAGAC 1846  
QY 7506 CCAAGCCTGTGTCTGTGAGTGAAGACATCTTCCAGACGAGTGTCAAGAACAG 7565  
DB 1847 CCAAGCCTGT 1906  
QY 7566 GCGGGGGGGGGGGGGGGGT 7625  
DB 1907 GCGGGGGAGCGG-----TGCAGGTCCCGGACAAAGGCAAGGAGGAGCGC 1948

OY	7626	TTCGACGAGCGGGGGCGCTGCACATCTCCGCCAACAAGAAAGAGCCCCTAGTCCGCT	7685
Db	1949	CAGCATGGAGCAGACAGCTGCTACTGACTCCCGAGA CTGAGAAAGAGACCCTTAGTCTGCT	2008
OY	7686	GCGCTCCATC-----CATCTGTCCGTCAAGACGGGCATCTTGCTG	7729
Db	2009	GCGGCTCATCTGCTCCGTCGCTCCACTCATCTGCTTCATCCAGAGCTGGCATTC--TGCTTG	2067
OY	7730	TCTTAAGCCTTAATTAAAGATCTCCGCGCCCGGGCGCTGGCCCTGTGAG----ACCTTACTCA	7785
Db	2068	TCTTAAGCCTTAATTAAAGATCTCCACCCCGGCGCTGGCCCTGTGAGCTTACACTCA	2127
OY	7786	GGGAGATGTTTACTTGCTGCTCGGGAAAGGAGGGGAAGGGCGG-----	7829
Db	2128	GGGAGATGTTTACTTGCTGCTCGAGAGGGGAGTGGACAGAGAGGGAGGGACAGCGGCGAC	2187
OY	7830	--GGAGGGGAGCAGCGGAGCGCTGTGGACGCA CACAGGCGGGCCAGGGCGGACAGGAC	7887
Db	2188	GAGGGGGGGGAGCAGCAATCGTGTEGTCAGTGGCACTCGTGACATGGGGGACAG--CGGGAC	2246
OY	7888	CCAAAAGCAGATGACACGCACTCCACGCGCATGGCC- -CCCCGAAATGATTTGGAAC	7945
Db	2247	CCACACACAGSCTGACAGGACCTTCATGCAACGCGCTCGCCCTTACCGCATTTGGAC	2306
OY	7946	CAAAGTCTAAACTAGAGCTCGAGCCCCCGGCGCCCTCTCGGCTCCCATCCGCTTACG	8005
Db	2307	CAAAGTCTAAACTGAATCTCGCGGTGATCTGCCCCCTCTGCTGCCCCAGCCGCTTG	2366
OY	8006	GCTCTGCAGAGATGGACGCGAGCCCTGTGCAGCCCCAGTGGCGCTGTTCCGGTCCAC	8065
Db	2367	GCTCTGCAGACAGACAGCTCTCCAGCTTATCTGCTCTTAATGCTGTCAATCGCAGTCT	2426
OY	8066	AGACTGCCCCAGCCAAGAGATTTGTGGAAACAAGTACGGCCAGTGGGGGACAAAG	8125
Db	2427	CCAAAGCTCACCCAGCCCAAGATGGGAGCCA-TAGACACAGTGTGGTGAC-AAAG	2486
OY	8126	GGCCAGGTGCGGCTGTGGGGGAGCGAAGCTCCGAGGACTGTGTTTTTTCACAT	8185
Db	2485	GGCCAGGTGCTGGCT- -GGGACAGGAATGTTGACAGGAATGGACCGTTTCCCGGCTGT	2542
OY	8186	CGTTGCCGACGCGGTGGGAGAGAAAGCAGATGTTAATGATGTGTGTTTACAGGGTAT	8245
Db	2543	TGCTGTGGCAACGGG---AGGAAAGGACAGCTGTAAAG--GTGTGGCTTACAGGGTAT	2596
OY	8246	ATTTTGTATCTTCAATGAATTAATTAGANGTTTTTACGGAAGGAAGACTTACCCAGT	8305
Db	2597	ATTTTGTATCTTCAATGAATTAATTAGACGTTCACACAAAGGAAGACTCGCCC-GT	2655
OY	8306	ATTATGCTGTGCTGCTTTTGTATCTCTGCTTACCGTTCAAGAGGCGTGTGACGGCCACA	8365
Db	2656	GTTTCTGCGCGCTGTGC--TTTGATCTTACTTACTTGTTCAGAGGACGTGCAAGC--AA	2712
OY	8366	GTCGGTACCCCATCATCTGCGAGACCAAGGGGGCGGGGACTGCTGTCACAGCCCGCTG	8425
Db	2713	GGCGGTGGCCCCACATACGTACGAGGACTTGGGGGTTCAGGGGCGTCCGGCACAGGCACTGTG	2772
OY	8426	-TGTCTCCCTCCCTCCCTCTCTTGGGAGAGATGAATTCAGATGGCTATTTCTGTGGCGGC	8484
Db	2773	CCCTTCCCCCACCCTTCTCTTGGGCAAAATGCACTTGATGCGTATTTCTGTGGCGCT	2832
OY	8485	ATTTCGCGAGGGTGGTATTTCTGTCAATTTTACACAGCTGTTCTAATTTAAAAAGCAT	8544
Db	2833	CTCTGTGCA---CGGCGGCAATCTGTCAATTATACATGTTGTTCCAAATTAAAGCAAT	2889
OY	8545	TATACTCCAAAAAAAAAAAAAAAAAAAA 8571	
Db	2890	ATACTCAAGTACCAACGTTAAAAAAA 2916	

RESULT 14			
AC073916/C			
LOCUS			
AC073916	205283 bp	DNA	linear
		PRI	27-MAR-2003

DEFINITION	Homo sapiens 12 BAC RP11-408I18 (Roswell) Park Cancer Institute Human BAC Library) complete sequence.
ACCESSION	AC073916
VERSION	AC073916.41
KEYWORDS	GI:29293998
SOURCE	HTG.
ORGANISM	Homo sapiens (human)
ORGANISM	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
REFERENCE	1 (bases 1 to 205283)
AUTHORS	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,

REFERENCE	Alsbrooks, S.L., Amarantunga, H.C., Ate, J.R., Ayele, M., Banks, T., Barbieri, J., Benton, J., Blinage, K., Blankenburg, K., Bonini, D., Bouck, J., Bowse, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, S., Christopher, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Dem, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emerling, S., Escotto, M., Falier, T., Ferriguto, D., Flagg, A., Ford, J., Foster, P., Franz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garcia, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hultky, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Joliver, S., Joudan, S., Karlsson, E., Kelly, S., Khan, U., King, L., Koeh, J., Kovar, C., Kretovc, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lightarge, O., Liu, C., Liu, J., Liu, W., Loubeleged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Mercher, S., Metzker, M., Miller, A., Miner, G., Milner, Z., Mitchell, T., Mohabadi, K., Montgometry, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkenwo, S., Ogun, M., Okwomu, G., Ogunyemi, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Plinius, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojudoakan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shoostari, N., Sisson, I., Sutogyan, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerias, A., Tamerias, K., Tang, H., Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umanai, K., Vaquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wellington, S., Williams, G., Williamson, A., Wleczky, R., Wooten, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zortilla, S., Kucheraipatti, R., Weinstein, G. and Gibbs, R.
TITLE	Unpublished
JOURNAL	Direct Submission
REFERENCE	2 (bases 1 to 205283)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (05-JUN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 205283)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (31-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	4 (bases 1 to 205283)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (01-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

5 (bases 1 to 205283)  
Worley, K. C.  
Direct Submission  
Submitted (06-MAR-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
6 (bases 1 to 205283)  
Worley, K. C.  
Direct Submission  
Submitted (15-MAR-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
7 (bases 1 to 205283)  
Worley, K. C.  
Direct Submission  
Submitted (27-MAR-2003) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Mar 27, 2003 this sequence version replaced gi:22038291.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

**CLONE LENGTH:** This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

### ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

**QUALITY OF INDIVIDUAL BASES:** this sequence meets stringent quality standards – estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES	Location/Qualifiers
source	1. 205283

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-408118"
/complement(1..116)
/rpc_family="MLT1J2"
/complement(185..545)
/rpc_family="MLT1A0"
STS
482..856
/standard_name="D12S820"
repeat_region
548..569
/rpc_family="AT_rich"
repeat_region
710..845
/rpc_family="(TCCA)n"

```

	repeat_region	complement(846..980)	
	/rpt_family="L2"		
	repeat_region	1320..1681	
	/rpt_family="L2"		
	repeat_region	2124..2158	
	/rpt_family="AT_rich"		
	repeat_region	2327..2353	
	/rpt_family="(T)n"		
	repeat_region	complement(2835..2914)	
	/rpt_family="MERS8B"		
	repeat_region	3015..3377	
	/rpt_family="L1MC4"		
	repeat_region	complement(3384..3882)	
	/rpt_family="LTR40a"		
	repeat_region	3908..3944	
	/rpt_family="U2"		
	repeat_region	3983..4182	
	/rpt_family="L1MC4"		
	repeat_region	4217..4256	
	/rpt_family="(GA)n"		
	repeat_region	5472..5647	
	/rpt_family="(TG)n"		
	repeat_region	5698..5870	
	/rpt_family="(TG)n"		
	repeat_region	complement(6009..6188)	
	/rpt_family="MIR"		
	repeat_region	6410..6484	
	/rpt_family="MIRB"		
	repeat_region	complement(6501..6815)	
	/rpt_family="A1USX"		
	repeat_region	7401..7603	
	/rpt_family="L2"		
	repeat_region	8096..8186	
	/rpt_family="MER96B"		
	repeat_region	complement(8187..8485)	
	/rpt_family="A1USX"		
	repeat_region	8486..8515	
	/rpt_family="MER96B"		
	repeat_region	9052..9174	
	/rpt_family="L2"		
	repeat_region	10287..10476	
	/rpt_family="(CGGG)n"		
	repeat_region	10415..10551	
	/rpt_family="G-rich"		
	repeat_region	10602..10631	
	/rpt_family="GC_rich"		
	repeat_region	10658..10684	
	/rpt_family="GC_rich"		
	repeat_region	10781..10801	
	/rpt_family="(CCCCAG)n"		
	repeat_region	10835..10859	
	/rpt_family="GC_rich"		
	repeat_region	11332..11465	
Query Match	13.2%;	Score 1132.4;	DB 8; Length 205283;
Best Local Similarity	97.7%;	Pred. No. 4.2e-138;	
Matches 1166; Conservative	3;	Mismatches 11;	Indels 13; Gaps 2;
OY	7368	CCGAGGTTCCAGCGCATTTCCCTACACCCCTGATCATCGGCTGCAGCGGGGTGCA	7427
Db	40848	CCGAGGTTCCAGCGCATTTCCCTACACCCCTGATCATCGGCTGCAGCGGGGTGCA	407899
OY	7428	TGGCTTCCCAACCCCAACCGGGGCTCCCGGGGAGCGGGGCCCTCGCTGGCCCCCAC	7487
Db	40788	TGGCTTCCCAACCCCAACCGGGGCTCCCGGGGAGCGGGGCCCTCGCTGGCCCCCAC	407299
OY	7488	ACCGCTGGAGCGAGAGCCCAAGCCACTCTCTGCTCGAGTACGAGACACTCTCCGACA	7547
Db	40728	ACCGCTGGAGCGAGAGCCCAAGCCACTCTCTGCTCGAGTACGAGACACTCTCCGACA	406659
OY	7548	GCGAGTACTCAGAACAGGCGGGGGGGGGCGGGCGGGCGGTGCTGCTGGCTCCAG	7607
Db	40668	GCGAGTACTCAGAACAGGCGGGGGGGGGCGGGCGGGCGGTGCTGCTGGCTCCAG	406211





3523	5368:	contig of 1456 bp in length
3529	5468:	gap of 100 bp
5469	6519:	contig of 1051 bp in length
6520	6619:	gap of 100 bp
6620	7852:	contig of 1233 bp in length
7853	7952:	gap of 100 bp
7953	8464:	contig of 512 bp in length
8465	8564:	gap of 100 bp
8565	10347:	contig of 1783 bp in length
10348	10447:	gap of 100 bp
10448	12145:	contig of 1698 bp in length
12146	13245:	gap of 100 bp
13246	14006:	contig of 1761 bp in length
14007	14106:	gap of 100 bp
14107	15358:	contig of 1252 bp in length
15359	15458:	gap of 100 bp
15459	17152:	contig of 1694 bp in length
17153	17252:	gap of 100 bp
17253	18800:	contig of 1548 bp in length
18801	18900:	gap of 100 bp
18901	20291:	contig of 1391 bp in length
20292	20391:	gap of 100 bp
20392	21889:	contig of 1498 bp in length
21890	21989:	gap of 100 bp
21990	24355:	contig of 2366 bp in length
24356	24455:	gap of 100 bp
24456	26757:	contig of 2302 bp in length
26758	26857:	gap of 100 bp
26858	28317:	contig of 1460 bp in length
28318	30781:	gap of 100 bp
30782	30881:	gap of 100 bp
30882	33789:	contig of 1908 bp in length
32790	33889:	gap of 100 bp
32890	35214:	contig of 2325 bp in length
35215	35314:	gap of 100 bp
35315	37181:	contig of 1867 bp in length
37182	37281:	gap of 100 bp
37282	39642:	contig of 2361 bp in length
39643	39742:	gap of 100 bp
39743	41978:	contig of 2236 bp in length
41979	42078:	gap of 100 bp
42079	44963:	contig of 2885 bp in length
44964	45063:	gap of 100 bp
45064	47265:	contig of 2202 bp in length
47266	47365:	gap of 100 bp
47366	49847:	contig of 2482 bp in length
49848	49947:	gap of 100 bp
49948	52474:	contig of 2527 bp in length
52475	52574:	gap of 100 bp
52575	54767:	contig of 2193 bp in length
54768	54867:	gap of 100 bp
54868	57938:	contig of 3071 bp in length
57939	58038:	gap of 100 bp
58039	60538:	contig of 2500 bp in length
60539	60638:	gap of 100 bp
60639	63422:	contig of 2784 bp in length
63423	63522:	gap of 100 bp
63523	66038:	contig of 2516 bp in length
66039	66138:	gap of 100 bp
66139	69185:	contig of 3047 bp in length
69186	69285:	gap of 100 bp
69286	72321:	contig of 3036 bp in length
72322	72421:	gap of 100 bp
72422	75240:	contig of 2819 bp in length
75241	75340:	gap of 100 bp
75341	77598:	contig of 2258 bp in length
77599	77698:	gap of 100 bp
77699	82485:	contig of 4787 bp in length
82486	82585:	gap of 100 bp
82586	86177:	contig of 3592 bp in length
86178	86278:	gap of 100 bp
86278	91129:	contig of 4852 bp in length

	* 91130	91239: gap of 100 bp	*
	91230	94475: contig of 3246 bp in length	*
	94476	94575: gap of 100 bp	*
	94576	98382: contig of 3807 bp in length	*
	98383	98482: gap of 100 bp	*
	98483	103891: contig of 5409 bp in length	*
	103892	103991: gap of 100 bp	*
	103992	109749: contig of 5758 bp in length	*
	109750	109849: gap of 100 bp	*
	109850	116075: contig of 6226 bp in length	*
	116076	116175: gap of 100 bp	*
	116176	122437: contig of 6262 bp in length	*
	122438	122537: gap of 100 bp	*
	122538	132392: contig of 9855 bp in length	*
	132393	132492: gap of 100 bp	*
	132493	141881: contig of 9389 bp in length	*
	141882	141981: gap of 100 bp	*
	141982	161970: contig of 19969 bp in length.	*
FEATURES	Location/Qualifiers		
source	1..161970		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/chromosome="12"		
	/map="12"		
	/clone="RP11-665C13"		
	/clone_1fb="RPCI-11 Human Male BAC"		
misc_feature	1..1052		
	/note="assembly_fragment"		
gap	1053..1152		
	/estimated_length=100		
misc_feature	1153..2378		
	/note="assembly_fragment"		
gap	2379..2478		
	/estimated_length=100		
misc_feature	2479..3812		
	/note="assembly_fragment"		
gap	3813..3912		
	/estimated_length=100		
misc_feature	3913..5368		
	/note="assembly_fragment"		
gap	5369..5468		
	/estimated_length=100		
misc_feature	5469..6519		
	/note="assembly_fragment"		
gap	6520..6619		
	/estimated_length=100		
misc_feature	6620..7852		
	/note="assembly_fragment"		
Query Match	11.1%; Score 948.4; DB 14; Length 161970;		
Best Local Similarity	97.6%; Pred. No.3,7e-114;		
Matches 965; Conservative	4; Mismatches 19; Indels 1; Gaps 1;		
OY	7574 GGCGCGGGGCGGTCTGTCAGTCCAGTGCCAGCAAGCAAGAAAGGCCCTTGCAGA	7633	
Dd	60639 GCGCGGGGGGGGGGGGGGGGCGAGTGTCAGATCCAGCGAAGCAAGGAACGGCCCTGCAGGA	60639B	
OY	7634 GCGGGGCGGCTGGCGCATCCCCAACCAAGAAAGAGGCCCGTAGCGGCGGCTCC	7693	
Dd	60699 GCAGGGGGGTTGGTCGACTCCCCAACCAAGAAAGAGGCCCTGAGTCGCGCTCC	60758B	
OY	7694 ATCCATCTGTCCTGCAGAGCGCGGATCTTGCTGTCTTAAGACCCTTAATAAGATTCCC	7753	
Dd	60759 ATCCATCTGTCCTGCAGAGCGCGGATCTTGCTGTCTTAAGACCCTTAATAAGATTCCC	60818B	
OY	7754 GCCCGGGGCTGGGCTGTGCAGACTTACTCAGGGAGATTAACTGTGTCTCGGGAGG	7813	
Dd	60819 GCCCGGGGCTGGGCTGTGCAGACTTACTCAGGGAGATTAACTGTGTCTCGGGAGG	60878B	
OY	7814 GAAGGGAAAGGGGCGGGGGAAGGGGGCACCGGACGGCTGTGGCAGCCACAACAAGCGGCCA	7873	
Dd	60879 GAAGGGAAAGGGGCGGGGGAAGGGGGCACCGGACGGCTGTGGCAGCCACAACAAGCGGCCA	60938B	



**THIS PAGE BLANK (USPTO)**